Landscape factors shaping host connectivity and pathogen dynamics in urban bobcats (*Lynx rufus*)

Christopher Peter Kozakiewicz

Bachelor of Science with First Class Honours

November 2018

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

School of Natural Sciences, University of Tasmania

UNIVERSITY*of* TASMANIA

AUSTRALIA

Thesis abstract

As a major factor affecting connectivity in wildlife populations, landscape heterogeneity can have substantial impacts on disease transmission. Urban development is a particularly acute contributor to landscape heterogeneity and declines in wildlife population connectivity, but few studies examine disease transmission and spread in urban environments. In this thesis, I utilise fine-scale genetic inference of how landscape heterogeneity shapes host connectivity and pathogen transmission in an urban environment. I focus on populations of bobcats, which are sensitive to urban development and indicators of connectivity in urban landscapes, from coastal southern California, one of the most urbanised landscapes in North America. Within these populations, I study feline immunodeficiency virus (FIV_{Lru}), a bobcat specific, rapidly mutating retrovirus and model for genetic inference of pathogen transmission in heterogenous landscapes.

The field of landscape genetics investigates how landscape heterogeneity affects genetic variation and has potential as a framework for studying pathogen transmission and spread. I conducted a comprehensive review of landscape genetic studies of pathogen dynamics (Chapter 2). I found that landscape genetics has been underutilised in disease ecology, partly due to a lack of cross-disciplinary awareness within the field of disease ecology, and a lack of landscape genetic frameworks for pathogen systems. I emphasise the utility of landscape genetics for disease ecology and highlight emerging frontiers, including recent phylogeographic approaches and multi-species analytical frameworks.

In Chapter 3, I conducted a landscape genomic study to investigate how urban and non-urban landscape factors are influencing gene flow among bobcats in southern California. I identified five genetically distinct populations, separated by major highways and urban development. Replicating landscape resistance analyses among these populations enabled me

to assess the generality of landscape effects on gene flow. I found that urbanisation had a pervasive impact on connectivity, influencing region-wide patterns of gene flow as well as locally within three populations, but that availability of riparian habitat may mitigate these urban impacts. This work demonstrates the value in replicating landscape genetic analyses across populations, showing that factors affecting connectivity in urbanising environments may vary depending on spatial scale and local landscape structure.

Having identified host population structure that is driven by major highways, I then (Chapter 4) conducted a phylogeographic analysis of FIV_{Lru} to investigate the impact of these barriers on FIV_{Lru}. Estimates of FIV_{Lru} divergence times dating back 118 years revealed a history of changes in pathogen transmission as urbanisation has increased. I identified strong phylogeographic structure, reflecting host population structure, and low FIV_{Lru} genetic diversity northwest of Los Angeles, suggesting relative isolation of FIV_{Lru} populations. Southeast of Los Angeles, FIV_{Lru} was genetically diverse with deep phylogenetic branches but little phylogeographic structure, suggesting divergence from host population structure. However, when accounting for variation in branch depth, movement rates of FIV_{Lru} across highways did not differ among populations. Divergence from host population structure southeast of Los Angeles was potentially a product of incomplete lineage sorting due to greater FIV_{Lru} sequence diversity and population size.

Lastly (Chapter 5), I implemented ecological phylogenetic tools to identify landscape and host factors influencing patterns of FIV_{Lru} phylogenetic differentiation and dispersal rates. Overall, I found that urbanisation plays less of a role in reducing FIV_{Lru} connectivity than for host connectivity. However, FIV_{Lru} sequences from bobcat capture locations that were more divergent in the amount of vegetation land cover were more distantly related. Specifically, this was the case for forest land cover northwest of Los Angeles and scrub land cover southeast of Los Angeles. My results suggest FIV_{Lru} transmission differs between areas

of high natural vegetation and areas of low natural vegetation (which are often urban areas). Further supporting the importance of vegetation for FIV_{Lru} transmission, I found a weak positive overall effect of vegetation density on FIV_{Lru} dispersal velocities.

In summary, this thesis: i) identifies and implements a variety of emerging methods for elucidating landscape effects on host and pathogen spatial genetic structure; ii) identifies factors affecting bobcat connectivity in a highly urbanised environment; iii) indicates how host population structure and landscape heterogeneity shape FIV_{Lru} phylogenetic structure and transmission dynamics; and iv) demonstrates the utility of replicating analyses of genetic structure across multiple populations and spatial scales to contextualise observed patterns and relationships. Collectively, this work represents a rare example of integrating genetic estimates of both host and pathogen connectivity in a heterogeneous landscape. These insights provide valuable information for managing an urban wildlife host-pathogen system, while showcasing the utility of landscape genetics and emerging ecological phylogenetic tools for studying connectivity in heterogeneous landscapes.

Acknowledgements

First and foremost, I would like to thank my supervisors, Scott Carver and Chris Burridge. I could not have hoped for two more supportive, hard-working, and capable mentors. It must have been disconcerting to have a student on the opposite side of the world for so long, and I appreciate not only your patience and trust in allowing me to spend that time, but also your attention and commitment to meeting regularly and keeping me on track when I needed it. I have always appreciated your accessibility and the efforts you each would go to in providing timely and always helpful feedback. This PhD has been a life-changing and overwhelmingly positive experience and I will forever be grateful to you for that.

To my unofficial supervisor-away-from-home, Chris Funk, thank you very much for welcoming me into your lab. You provided not just the facilities and expertise for me to learn and conduct my work, but a stimulating and enriching environment through which I could broaden my horizons and grow as a scientist. To you and Kim Hoke and the combined Funk-Hoke labs, I had a wonderful couple of years with you all, and to work alongside you and share in your weekly lab meetings was such a rewarding experience.

My PhD formed but one small part of the broader *FELIDAE* project. This is an academically brilliant yet wonderfully down-to-Earth and fun group of people, and our annual meetings at the CSU Mountain Campus were always an exciting and enjoyable time. My thanks and congratulations go to the PIs who conceived and led the project: Sue VandeWoude, Kevin Crooks, Chris Funk, Scott Carver, Meggan Craft, and Holly Ernest. Sue, you also have my thanks for providing me the opportunity to work in your lab.

To the *FELIDAE* postdocs: Simona Kraberger, Patricia Salerno, Daryl Trumbo, Erick Gagne, and Nick Fountain-Jones. I learnt a lot of new skills during my PhD, and many of those were passed down from you. Thank you all for your expertise and your time – I could

not have done it without you. Not only that, you each have also been excellent role models as postdocs and I hope to follow your examples as I move into that stage of my career. Thanks also to Jennifer Malmberg and Justin Lee for all the helpful advice when I was trying to get the FIV sequencing to work.

Each one of the nearly 300 samples I used in my work was carefully collected from a fast, intelligent, secretive, and I assume uncooperative animal with sharp claws and teeth! Catching them is no small task and to do it so many times required a huge effort over many years by many dedicated people: Erin Boydston, Lisa Lyren, Kevin Crooks, Seth Riley, Laurel Serieys, Megan Jennings, and the many technicians and others who presumably worked with them in the field over the years. My work would have been impossible without these people. Thanks Erin, Megan, and Seth in particular for kindly showing Nick and I around your field sites – this first-hand experience of the landscape was invaluable. I also wish to acknowledge the support of the National Science Foundation for my research, and the Australian Government Research Training Program for my scholarship.

I would like to generally thank the University of Tasmania Division of Biological Sciences. This is a wonderful community of people and such a friendly and supportive environment to work in. There are many of you and I cannot name everyone, but here are just a few. I want to thank the admin staff, specifically Clancy Carver, Jodi Noble, and Morgan Green, for always making things happen. Thank you to my office mates: Glen Bain, Mercy Ndalila, Angie Reid, Hannah Cliff, Tamieka Fraser, and David Lizarraga, for the daily conversation, encouragement, and shared hardship. Thank you to the Funky Genomics Group (Chris still doesn't know we named our genomics reading group after him after he attended our inaugural meeting), especially the regulars: Floriaan Devloo-Delva, Mingxin Liu, Kirstin Proft, and Alynn Martin, for the enlightening discussions. Again, thank you Kirstin for the many many discussions about landscape genetics and related topics, cups of tea, and

encouraging words, and also to Alynn for the regular chats about BEAST and everything else under the sun! Thanks Riana Gardiner for being a wonderful housemate and great friend.

I especially want to thank my parents, Della and Richard Kozakiewicz, for always being supportive, for putting up with me living at home at the age of 30, and for being there when I've needed it. Cheers to Rudy Fidao and Tristan Sell for the beers. Finally, I want to acknowledge my partner, Alisha Shah. I mentioned above that my PhD was a life-changing experience. There are many reasons for that, but, Alisha, you are one of the foremost. Although the distance has been hard, our conversations from opposite sides of the world are a daily highlight. Thank you for your love and constant support, and I truly look forward to our future together.

Statements and declarations

Declaration of originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis and, to the best of my knowledge and belief, no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Signed: _____ Date: 20/11/2018

Authority of access

This thesis may be made available for loan. Copying and communication of any part of this thesis is prohibited for two years from the date this statement was signed; after that time limited copying and communication is permitted in accordance with the Copyright Act 1968.

Signed: _____ Date: 20/11/2018

Statement regarding published work contained in thesis

The publisher of the paper comprising Chapters 2 holds the copyright for that content and access to the material should be sought from the respective journal. The remaining non-published content of the thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Due to the inclusion of published and soon-to-be published material, there is unavoidable repetition of material between chapters in this thesis.

Signed: _____ Date: 20/11/2018

Statement of co-authorship

The following people and institutions contributed to the publication of work undertaken as part of this thesis:

Christopher P. Kozakiewicz, School of Natural Sciences, Candidate

Scott Carver, University of Tasmania, Supervisor

Christopher P. Burridge, University of Tasmania, Supervisor

W. Chris Funk, Colorado State University

Sue VandeWoude, Colorado State University

Kevin R. Crooks, Colorado State University

Meggan E. Craft, University of Minnesota

Holly B. Ernest, University of Wyoming

Nicholas M. Fountain-Jones, University of Minnesota

Patricia E. Salerno, Colorado State University

Daryl R. Trumbo, Colorado State University

Roderick B. Gagne, Colorado State University

Erin E. Boydston, United States Geological Survey

Lisa M. Lyren, United States Geological Survey

Robert N. Fisher, United States Geological Survey

Megan K. Jennings, San Diego State University

Laurel E.K. Serieys, University of California

Seth P.D. Riley, California National Park Service

Justin S. Lee, Colorado State University

Simona J. Kraberger, Arizona State University

Chapter 2 was accepted for publication in the journal *Evolutionary Applications* in June 2018

Kozakiewicz CP, Burridge CP, Funk WC, VandeWoude S, Craft ME, Crooks KR, Ernest HB, Fountain-Jones NM, Carver S (2018) Pathogens in space: advancing understanding of pathogen dynamics and disease ecology through landscape genetics. *Evolutionary Applications*, **11**, 1763–1778

Author contributions: C.P.K. conducted this review, compiled literature, and wrote the manuscript with critical feedback from all authors. W.C.F., S.V., S.C., K.R.C., H.B.E., and M.E.C. conceived of the project that led to this review. S.C., C.P.B, N.M.F.J., and M.E.C. provided early input that contributed to the present framing and structure of this review, while S.V., and S.C. assisted in collating information from the literature search.

Chapter 3 is currently in review for publication in the journal *Molecular Ecology*

Kozakiewicz CP, Burridge, CP, Funk, WC, Salerno, PE, Trumbo, DR, Gagne, RB, Boydston, EE, Fisher, RN, Lyren, LM, Jennings, MK, Riley, SPD, Serieys, LEK, VandeWoude, S, Crooks, KR, Carver, S (*in review*) Urbanization reduces genetic connectivity in bobcats (*Lynx rufus*) at both intra- and inter-population spatial scales. *Molecular Ecology*

Author contributions: C.P.K. conducted the laboratory work, bioinformatics, statistical analysis, and wrote the manuscript with critical feedback from all authors. P.E.S., R.B.G., and D.R.T. provided input on the laboratory work, bioinformatics and statistical analyses. E.B., L.L., M.K.J., S.P.D.R., K.R.C., and L.E.K.S. collected samples. W.C.F., S.V., K.R.C., and S.C. conceived of the project. C.P.B., S.C., and W.C.F. oversaw the project and provided conceptual guidance.

Chapter 4 is currently in preparation for publication in the journal *BMC Evolutionary Biology*

Kozakiewicz CP, Burridge, CP, VandeWoude, S, Lee, JS, Kraberger, SJ, Fountain-Jones, NM, Funk, WC, Boydston, EE, Lyren, LM, Jennings, MK, Riley, SPD, Serieys, LEK, Crooks, KR, Carver, S (*in prep*) Does the virus cross the road? Viral phylogeographic patterns among urban bobcat populations. *BMC Evolutionary Biology*

Author contributions: C.P.K. conducted the laboratory work, statistical analysis, and wrote the manuscript with critical feedback from all authors. J.S.L. and S.J.K. provided guidance on the laboratory work. N.M.F.J. and C.P.B. provided guidance on the statistical analysis. E.E.B., L.M.L., M.K.J., S.P.D.R., K.R.C., and L.E.K.S. collected samples. W.C.F., S.V., K.R.C., and S.C. conceived of the project. C.P.B. and S.C. oversaw the project and provided conceptual guidance.

Chapter 5 is currently in preparation for publication in the journal *Virus Evolution*

Kozakiewicz CP, Burridge, CP, VandeWoude, S, Lee, JS, Kraberger, SJ, Fountain-Jones, NM, Funk, WC, Boydston, EE, Lyren, LM, Jennings, MK, Riley, SPD, Serieys, LEK, Crooks, KR, Carver, S (*in prep*) Virus phylogeography in urban bobcat populations reveals increased transmission in vegetated habitats. *Virus Evolution*

Author contributions: C.P.K. conducted the laboratory work, statistical analysis, and wrote the manuscript with critical feedback from all authors. J.S.L. and S.J.K. provided guidance on the laboratory work. N.M.F.J. provided guidance on the statistical analysis. E.E.B., L.M.L., M.K.J., S.P.D.R., K.R.C., and L.E.K.S. collected samples. W.C.F., S.V., K.R.C., and S.C. conceived of the project. C.P.B. and S.C. oversaw the project and provided conceptual guidance.

We the undersigned agree with the above stated “author contributions” for each of the above published (or submitted) peer-reviewed manuscripts contributing to this thesis:

Signed: _____ _____

Dr. Scott Carver
Supervisor
School of Natural Sciences
University of Tasmania

Dr. Christopher Burridge
Supervisor
School of Natural Sciences
University of Tasmania

Date: 20/11/2018 20/11/2018

Signed: _____

Assoc. Prof. Greg Jordan
Head of Discipline of Biological Sciences
School of Natural Sciences
University of Tasmania

Date: 20/11/2018

Table of contents

Chapter 1

General introduction

Landscapes are spatially heterogeneous, and this heterogeneity can have substantial impacts on wildlife, including changes in species diversity, interspecific interactions, species demographic structure, population dynamics, population genetics, and pathogen exposure and transmission (Pickett & Cadenasso, 1995). A primary process by which these impacts occur is through variation in functional connectivity, defined as the degree to which the landscape facilitates or impedes wildlife movement among habitat patches (Taylor et al., 1993). Thus, understanding how heterogeneity in specific landscape factors influence functional connectivity for wildlife can provide insights into many ecological processes, including pathogen transmission.

Anthropogenic land use is a dominant contemporary contributor to landscape heterogeneity (Fahrig et al., 2011). Urban development is a particularly acute example of this, and its impacts on wildlife connectivity are widely acknowledged and studied (Bierwagen, 2007). It is broadly understood that urbanisation can also have impacts on pathogen prevalence and dynamics (Brearley et al., 2013; Becker et al., 2015). However, host movement and pathogen spread are not always proportional. Indeed, a meta-analysis by Mazé-Guilmo *et al.* (2016) demonstrated that the genetic structure (a product of variation in connectivity) of parasites is frequently decoupled from that of hosts. Yet, research examining both host and pathogen connectivity in the context of landscape heterogeneity is rare, and few urban studies directly examine the processes of transmission and spread themselves. Thus, the mechanisms driving observed responses of prevalence to urbanisation are often poorly understood.

A major reason for the paucity of research linking pathogen transmission to landscape heterogeneity is that pinpointing the exact locations of transmission events, such that researchers can relate these to landscape factors, is difficult. Transmission events themselves are almost impossible to observe, so contacts among individuals are frequently used as a proxy (Craft & Caillaud, 2011). Contacts may in rare cases be observed directly, such as in African lions (Craft et al., 2009), but are typically inferred using spatial and/or temporal overlap; determined using techniques such as mark-recapture (Perkins et al., 2009), motion-activated cameras (Atwood et al., 2009), telemetry (Leu et al., 2010), or proximity loggers (Hamede et al., 2009). These methods are logistically challenging to employ, and lack precision because it's unclear whether a given contact was among infected individuals or was even sufficient for transmission. Instead, genetic tools are emerging that enable spatial genetic structure of pathogens to be assessed to infer contemporary transmission histories at sufficiently fine spatial and temporal scales commensurate with those of heterogeneous landscape factors (Archie et al., 2009; Gilbertson et al., 2018).

Model systems are valuable for evaluating emerging methodologies and testing broad ecological questions. The empirical work I present in this thesis focuses on bobcats (*Lynx rufus*) in coastal southern California and a virus they transmit intraspecifically, feline immunodeficiency virus (FIV_{Lru}). I implement robust and novel genetic approaches to illuminate the relationships among host dispersal, pathogen dynamics, and landscape heterogeneity in an urbanising environment. Coastal southern California is one of the most urbanised regions in North America (U.S. Census Bureau, 2016), yet, among the dense urban sprawl, there remain relatively large areas of natural habitat that are becoming increasingly fragmented and isolated (Thomassen et al., 2018). Bobcats are solitary, territorial carnivores that rely on small mammalian prey, predominantly lagomorphs and rodents in urban areas (Riley et al., 2010). They are good indicators of wildlife connectivity in coastal southern

California, having some tolerance to anthropogenic disturbance but overall relying on large areas of natural habitat with high connectivity among patches to maintain adequate rates of gene flow (Crooks, 2002).

Bobcats carry various pathogens, including FIV_{Lru} (VandeWoude & Apetrei, 2006; Lagana et al., 2013; Carver et al., 2016). FIV is a retrovirus with a number of characteristics that make it a suitable model for studying pathogen dynamics in heterogeneous landscapes. It is directly transmitted and generally forms species-specific strains, meaning that transmission only occurs through contacts among conspecifics. FIV forms life-long infections and has little overt pathology, enabling sampling of proviral DNA at any time post-infection. Lastly, the FIV genome mutates very rapidly, producing substantial genetic variation through which recent transmission may be inferred (Biek et al., 2003, 2006; Lee et al., 2012; Fountain-Jones et al., 2017a, 2017b). Combined, the coastal southern California landscape, bobcat host, and FIV_{Lru} pathogen form a powerful system for examining host connectivity and pathogen dynamics in an urban environment. Long-term field studies by several collaborating organisations have accumulated almost 15 years' worth of blood and tissue samples (Lyren et al., 2006; Riley et al., 2006; Lyren et al., 2008a, 2008b; Jennings & Lewison, 2013; Serieys et al., 2015), adding further utility to this system for meeting the objectives of this thesis.

1.1 Thesis objectives and structure

The overarching objectives of this thesis are 1) to understand how pathogen transmission and spread is influenced by host connectivity, and 2) to understand how heterogeneous landscape factors shape both of these processes in an urban setting. To address these broad objectives, I present in the following chapters four core studies (each forming

independent publishable units), bounded by this general introduction and a general discussion. I outline each of these chapters below.

Landscape genetics aims to investigate how heterogeneous landscape factors influence genetic variation (Manel et al., 2003; Storfer et al., 2007; Manel & Holderegger, 2013). While this field has traditionally been applied to the study of free-living organisms, landscape genetic tools hold considerable potential for studying pathogens in heterogeneous landscapes (Biek & Real, 2010). Yet, landscape genetics remains relatively under-utilised in disease ecology. Chapter 2 presents a comprehensive review of how landscape genetics is being applied to better understand pathogen dynamics (Kozakiewicz et al., 2018a). In that chapter, I characterise the existing body of work at the interface of landscape genetics and disease ecology. I also examine established landscape genetic methods and their utility in disease ecology, detailing both their realised and potential applications. Finally, I identify emerging frontiers in the landscape genetic study of infectious diseases and discuss some novel frameworks for addressing challenges in this multidisciplinary field. This review makes a substantial contribution to the direction of landscape genetics within the field of disease ecology and provides an important background for the subsequent chapters of this thesis.

A detailed understanding of the factors affecting connectivity in host species provides valuable context for studies of pathogen dynamics. In Chapter 3, I present a landscape genomic study to assess how urban and non-urban landscape factors are influencing gene flow among bobcats in coastal southern California. An often-neglected consideration in landscape genetic research is to assess the generality of landscape genetic effects among populations and across varying spatial scales (Short Bull et al., 2011). The distribution of bobcat sampling among several populations separated by major highways enables me to employ a novel replicated landscape genomic design whereby the effects of different landscape factors on gene flow can be examined locally for each population, and then

generally at a regional scale. These populations vary in landscape composition and range size, allowing me to contextualise the landscape genetic relationships I observe with respect to these factors. Thus, this chapter demonstrates the utility of replication in landscape genetic studies and provides valuable insights into factors affecting bobcat connectivity in coastal southern California.

In Chapter 4, I present a phylogeographic study of FIV in coastal southern Californian bobcats, which investigates the impact of major anthropogenic barriers to host connectivity on pathogen phylogenetic structure (Kozakiewicz et al. *in prep*). Phylogeography has traditionally been used to study historical processes in free-living organisms (Wang, 2010), but the coincident evolutionary and ecological dynamics of many rapidly mutating pathogens enable phylogenetic inference of recent transmission among host populations and individuals (Grenfell et al., 2004; Ypma et al., 2013; Biek et al., 2015). By constructing a phylogeny incorporating sampling dates, which enables explicit estimation of divergence dates, I assess the degree of congruence among host and pathogen population genetic structure with respect to major roads and urban barriers. I also demonstrate how connectivity in coastal southern California has changed as urban development in this region has expanded over time.

In Chapter 5, I present an ecological phylodynamics study that draws upon the work in prior chapters to investigate how host and fine-scale landscape factors influence FIV dynamics in coastal southern California. Ecological phylodynamics is an emerging approach that incorporates methods from landscape genetics and viral phylodynamics (Grenfell et al., 2004; Volz et al., 2013) to understand how landscape and other ecological factors shape phylogenetically-inferred transmission networks (Fountain-Jones et al., 2017c). This approach holds great potential for elucidating pathogen dynamics but has rarely been implemented. I employ two complementary techniques to, firstly, test host and landscape factors influencing FIV gene flow, and secondly, test factors influencing rates of FIV lineage

spread across the landscape. By quantitatively linking pathogen phylogenetic relationships to both the landscape and the host, this chapter provides extensive insights into factors shaping transmission networks.

Finally, in Chapter 6, I synthesise my findings and discuss the broader contribution of this thesis to our understanding of host connectivity and pathogen dynamics in urban environments. I discuss potential management implications of this work for both the bobcat-FIV_{Lru} system specifically and for the management of urban wildlife diseases more broadly. I also propose several directions for future research that have been highlighted by my work. These include the implementation of emerging methodologies and the study of additional populations and systems, which have the capacity to both build upon my findings and advance new frontiers in urban wildlife disease research.

Chapter 2

Pathogens in space: advancing understanding of pathogen dynamics and disease ecology through landscape genetics

2.1 Abstract

Landscape genetics has provided many insights into how heterogeneous landscape features drive processes influencing spatial genetic variation in free-living organisms. This rapidly developing field has focused heavily on vertebrates, and expansion of this scope to the study of infectious diseases holds great potential for landscape geneticists and disease ecologists alike. The potential application of landscape genetics to infectious agents has garnered attention at formative stages in the development of landscape genetics, but systematic examination is lacking. We comprehensively review how landscape genetics is being used to better understand pathogen dynamics. We characterise the field and evaluate the types of questions addressed, approaches used, and systems studied. We also review the now established landscape genetic methods and their realised and potential applications to disease ecology. Lastly, we identify emerging frontiers in the landscape genetic study of infectious agents, including recent phylogeographic approaches and frameworks for studying complex multi-host and host-vector systems. Our review emphasises the expanding utility of landscape genetic methods available for elucidating key pathogen dynamics (particularly transmission and spread) and also how landscape genetic studies of pathogens can provide insight into host population dynamics. Through this review we convey how increasing awareness of the complementarity of landscape genetics and disease ecology among practitioners of each field promises to drive important cross-disciplinary advances.

2.2 Introduction

The field of landscape genetics seeks to identify relationships between heterogeneous landscape features and genetic variation in free-living organisms, and has become a popular method for investigating drivers of processes such as gene flow, genetic drift, and selection. (Manel et al., 2003; Manel & Holderegger, 2013). Landscape genetics has grown substantially since its formal inception in 2003, facilitated by technological advances that have increased the availability of molecular and landscape data in conjunction with more powerful computational and analytical approaches. Landscape genetics is fuelled by a steady stream of new ideas and methodologies, which, while exciting, can contribute to a lack of consensus or consistency in some key aspects. These aspects include the formulation of research questions, sampling strategies, analytical methods (Balkenhol et al., 2009b; Wagner & Fortin, 2013; Richardson et al., 2016) and even the identity of the field itself (Storfer et al., 2007; Dyer, 2015). In fact, landscape genetics has yet to develop its own comprehensive, unifying theory for linking spatial and temporal landscape heterogeneity to genetic variation (Balkenhol et al., 2016b). While these issues are expected to be remedied as the field matures, many suggestions have been made to facilitate this progress. These have included calls for an increase in cross-disciplinary collaboration (Balkenhol et al., 2009a) and an expansion of the scope of landscape genetic research beyond its current emphasis on vertebrates (Dyer, 2015; Balkenhol et al., 2016b) and, particularly, mammals (Kozakiewicz et al., 2018b).

One logical avenue for cross-disciplinary expansion of landscape genetics is in disease ecology (Biek & Real, 2010). Elucidating the specific influences of landscape features on pathogen transmission can provide key insights into the processes that affect disease risk and incidence. However, accomplishing this has been a challenge for disease ecologists (Ostfeld et al., 2005). Indeed, the field of spatial epidemiology has only recently

begun to emphasise the use of explicit landscape approaches in studies of spatial heterogeneity in infectious disease (i.e. “landscape epidemiology”; Ostfeld et al., 2005; Meentemeyer et al., 2012). A major challenge for the study of landscape epidemiology, a field which does not traditionally implement genetic approaches, is that it is typically dependent on the ability to identify the location and timing of transmission events such that they can be compared to landscape features of interest. Transmission events are essentially impossible to observe, so disease ecologists often assume that contacts between infected and susceptible individuals are a reasonable proxy for transmission. Such contacts generally must be inferred indirectly using methods such as proximity collars, mark-recapture, or telemetry, often using spatial overlap as a proxy for contact (Craft & Caillaud, 2011). These methods are logistically challenging to employ, and whether an inferred contact resulted in transmission is uncertain (Craft, 2015). Further, much landscape epidemiological research uses infection or exposure data to indicate past transmission, but these methods provide static snapshots of pathogen prevalence and may be inappropriate for inferring how transmission or spread has occurred (or is occurring) over time (Meentemeyer et al., 2012).

The spatial distribution and movement of hosts are major factors affecting the likelihood, timing, and spatial patterns of pathogen transmission and spread (Dougherty et al., 2018). Landscape genetics can identify landscape factors that are important drivers of host population structure. These landscape factors can determine the spatial configuration of a population, its density, its connectivity with other populations, its demographic structure, and its genetic health – all of which have implications for the dynamics of micro-organisms infecting the host species (Spielman et al., 2004; Ellis et al., 2010; Prentice et al., 2014). Further, pathogen dynamics can be inferred directly using pathogen genetic data (Archie et al., 2009; DeCandia et al., 2018) and incorporated into landscape genetic analyses. Understanding specifically how infectious agents respond to the influence of landscape

factors on hosts enables us to predict how such agents might spread based on present landscape configurations, as well as under potential future landscape scenarios (Real & Biek, 2007). This knowledge can subsequently inform management efforts at the population level (such as vaccination targeted at key regions, culling, etc.), as well as broader decisions relating to the management of the landscape itself, which is a key aim of landscape genetics generally (Segelbacher et al., 2010; Manel & Holderegger, 2013). Landscape genetics is being applied by managers at relatively low rates compared to related ecological fields such as landscape ecology, conservation biology, and telemetry research (Bowman et al., 2016). Therefore, studies that contribute to the management of disease agents within populations could increase the practical impacts of landscape genetics significantly. However, the conceptual underpinnings of pathogen landscape genetics are not fully developed, and the methodologies employed are diverse and potentially confusing for new practitioners.

Here we investigate how landscape genetic techniques are being used to better understand dynamics of micro-organisms infecting host species. In conducting this review, we aim to both advocate and facilitate landscape genetic research involving disease-causing organisms. We first evaluate the use of landscape genetics in disease ecology, including the types of questions addressed, the approaches used, and the infectious agents studied. We then review established landscape genetic methods and their realised and potential applications to disease ecology. Finally, we identify emerging frontiers in the landscape genetic study of pathogens that hold significant potential for advancing research in this field.

Landscape genetics was first implemented in the study of rabies virus by Real et al. (2005), offering an approach to overcome many feasibility issues associated with understanding landscape influences on pathogen transmission. The landscape genetic approach to studying disease was later reviewed by Biek and Real (2010), who were optimistic about its growth and future use. In particular, they noted that microparasites, such

as viruses, are well-suited to landscape genetic study due to their rapid mutation rate and potential spatial genetic structure that can be compared to heterogeneous landscape features at fine temporal and spatial scales. Analyses could be conducted using both pathogenic organisms as well as agents that do not cause significant diseases in their hosts (e.g., Biek et al., 2006). They also identified that methodologies such as GIS, which are commonly employed both in the wider landscape genetics literature and in spatial studies of infectious disease, had not been widely implemented in molecular epidemiology (Archie et al., 2009). Further, other popular landscape genetic tools, such as those focused on differential landscape permeability (e.g., least-cost paths), were greatly under-used despite compatibility with pathogen spatial genetic data.

Similar to landscape genetics, landscape epidemiology is an interdisciplinary field undergoing rapid development driven by technological advancements, and arguably still working to develop clear directions for future research (Meentemeyer et al., 2012). It is therefore likely that the interface of these two fields (i.e., where landscape genetics is used in epidemiology) is similarly challenged, perhaps to the extent that its potential is remaining unrealised. We thus believe it is timely to revisit the body of research that combines landscape genetics and landscape epidemiology, leveraging the work done both prior and subsequent to Biek and Real's (2010) earlier review into clear directions for future research.

2.3 Current applications of landscape genetics in disease ecology

2.3.1 Literature search

We conducted a literature search in February 2018 using the ISI Web of Science database with the following terms:

TS=(("landscape genetic" OR "landscape genom*") AND (disease* OR pathogen* OR parasit* OR virus* OR virol* OR epidem* OR infect* OR transmi*))*

The search returned 133 results. We read each article and retained the 51 empirical papers that used landscape genetic methods to address questions related to pathogens (see Appendix Table A2.1). We excluded reviews (n = 15), meeting abstracts (n = 1), purely methods-based papers (n = 6), and articles that identified as or mentioned landscape genetics but did not sufficiently incorporate landscape factors or genetic data into the study (n = 32), studies that referred to any of our pathogen-related search terms without it being a primary motivation for the study (n = 21), and studies that used words like "transmit" or "parasite" outside of the context of infectious agents (such as the transmission of behaviours) (n = 6). One paper was excluded due to a lack of access at our institutions. Studies that qualitatively discussed landscape with respect to genetic variation were kept, although one might argue that landscape genetics requires quantitative testing of landscape effects. We classified each paper according to the type of host system studied (plant, wild animal, domestic animal, human), the type of pathogen studied (bacterium, protozoan, virus, prion, fungus, macroparasite, transmissible cancer), and the source of genetic data (host, pathogen, vector), and we estimated the severity of disease that each studied pathogen causes in its sampled host or vector. We also categorised each article according to its general conceptual approach. Most examples described in this paper were found in our literature search, while several other examples were cited by papers from our search and subsequently also discussed here.

Following publication of the first study using landscape genetics to investigate disease in 2005, there was little further research in this area until 2009, which saw a rapid increase in the number of publications (Figure 2.1a). This increase coincided with two prominent review articles (Archie et al., 2009; Biek & Real, 2010) that were strong proponents of a landscape genetics approach to disease ecology and expressed optimism about its future use. The rate of publication has remained relatively steady (and arguably low) since then, with none of the subsequent seven years recording more publications than in 2009, when six papers were published. However, 10 articles using landscape genetics to investigate disease were published in 2017, potentially indicating increasing interest in this area of research.

A majority of studies (27 of 51) used genetic data from the host for comparison with landscape features (Figure 2.1b). This is likely because DNA is easier to obtain from larger, free-living hosts than for pathogens, and methods for genotyping and characterising host spatial genetic variation are more familiar to landscape geneticists, who predominantly study free-living organisms (Storfer et al., 2010). Among pathogens that are associated with a particular animal vector, the vector is often genotyped (9 of 14 studies of vector-borne diseases), as vectors such as ticks or mosquitos are also easily sampled, and vector gene flow can be used as a proxy for pathogen spread. Vectors can be targeted for population control as a means of limiting pathogen spread, which makes their study of immediate relevance to wildlife and livestock managers (Townson et al., 2005). Pathogen genetic data are used in only 16 of 51 pathogen landscape genetic studies, which was somewhat surprising considering that the pathogen is the primary motivation behind many of the reviewed studies. One study included both host and pathogen genetic data (Talbot et al., 2017).

Viruses were the most frequently studied type of infectious agent (14 of 51 studies; (Figure 2.1c). In general, viruses evolve more rapidly than other microparasites, which makes them well-suited to study of genetic variation for inference of transmission history (Grenfell

et al., 2004; Archie et al., 2009). However, a majority of landscape genetic studies involving viruses used host genetic data, potentially reflecting the relative difficulty of obtaining viral data, which we discuss later in this section. Instead, the high representation of viruses is largely due to the considerable effort devoted to studying rabies, which comprised half of all landscape genetic studies on viral systems. Rabies is one of the most well-known wildlife pathogens globally, due to its negative impacts on wildlife, domestic animal, and human health (Gordon et al., 2004). Large outbreaks have occurred in North American and European wildlife in recent years, where considerable resources have been devoted to its management (Holmala & Kauhala, 2006; Slate et al., 2009). Animals infected with rabies also often exhibit behavioural changes that may make them easier to identify (Lefèvre et al., 2009), potentially aiding sampling of infected individuals.

We broadly define three distinct conceptual approaches by which landscape genetics has been used to study infectious agents (Figure 2.1d). These are the prediction of agent spread using genetic information from the host or vector; the use of host or vector genetic information to explain existing spatial variation in infection risk or prevalence; and the use of genetic information from the infectious agent to directly study transmission and spread. The remainder of this section will address each of these approaches in turn.

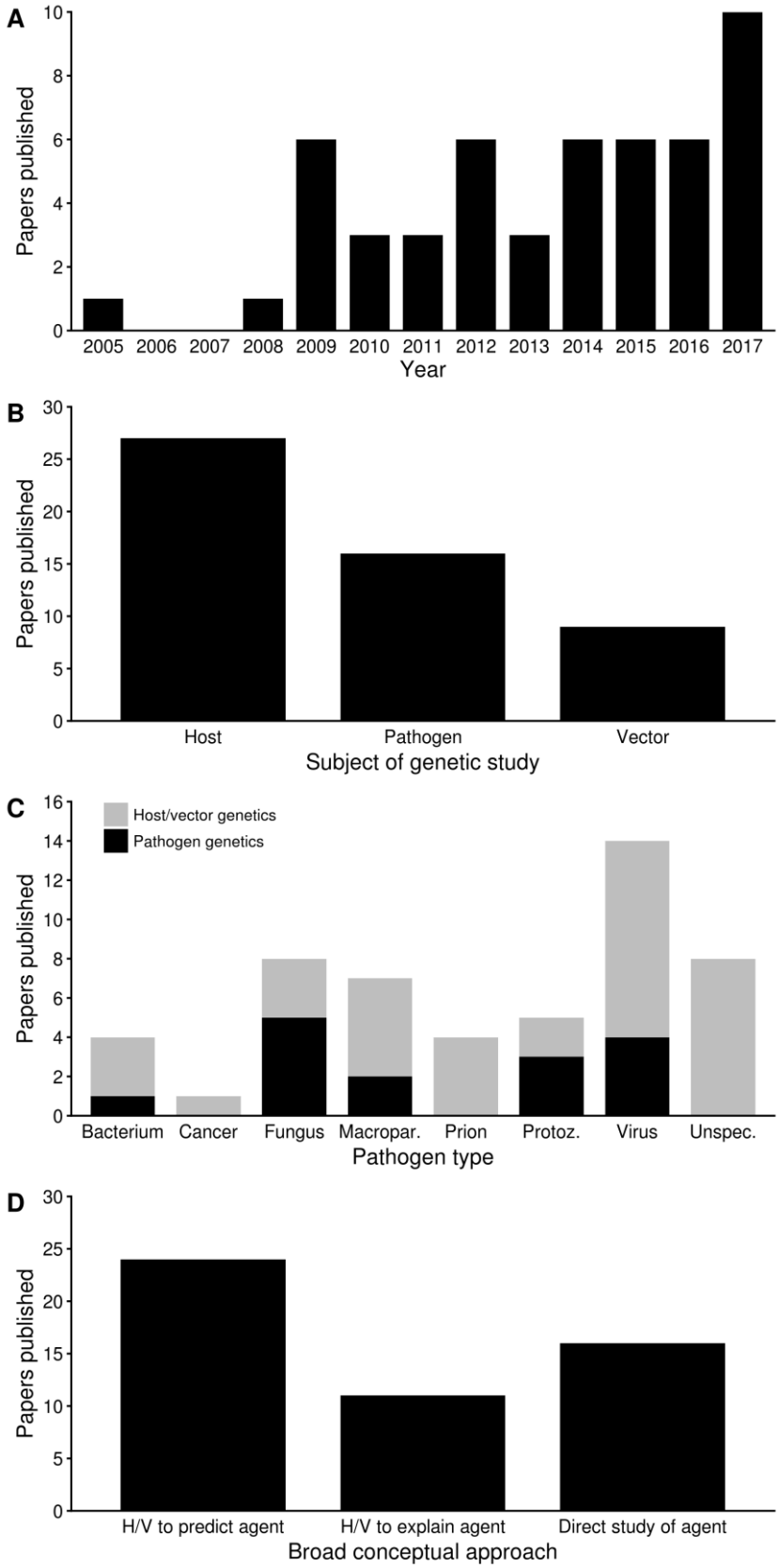

Figure 2.1. Papers using landscape genetic approaches for the study of infectious agents. a) Number of publications per year that met our search criteria. b) Number of publications using genetic data from each of the host, agent, or vector species. c) Number of publications studying pathogens by type, with genetic data source indicated for each type ("unspecified"

typically involves studies of a hypothetical agent or estimates of overall pathogen exposure, such as inferred by immune-linked loci). d) Number of publications adopting each of our broadly identified conceptual approaches for applying landscape genetics to the study of pathogens/infectious agents – using host/vector genetics to predict agent spread, using host/vector genetics to explain agent spread/distribution, and using pathogen genetics to directly study agent spread.

2.3.2 Host or vector genetic variation as a predictor of agent spread with respect to landscape

Because the spread of many microparasites (particularly directly transmitted forms) is facilitated by movement of free-living hosts or vectors, the risk of spread of the agent with respect to heterogeneous landscape features can be estimated by relating those features to host/vector gene flow. This approach represents a direct application of the conventional landscape genetic paradigm to the study of disease transmission, where a typical animal landscape genetic study is interpreted in the context of the pathogenic organism. This can provide useful indications of the potential for individuals carrying pathogens to disperse across particular landscape features, which can be used to inform management efforts. For example, DeYoung et al. (2009) identified long-distance gene flow among gray fox populations in Texas that was unrelated to landscape features tested, determining that current rabies oral vaccination plans should be expanded given the high potential for long-distance host movement. In another rabies study, landscape genetics was used to characterise striped skunk dispersal across riverine and highway barriers to assess their utility as barriers to pathogen spread (Talbot et al., 2012).

Using host or vector genetic data to predict pathogen spread is attractive as it avoids sampling of the agent itself, which may be substantially more difficult, especially in wildlife populations. Identification of infected hosts often requires laboratory testing, and may require specific, potentially invasive sampling approaches (e.g., necropsy) for accurate diagnosis.

Additionally, extensive sampling may be required to obtain adequate sample sizes when prevalence is low and must be conducted strategically to capture spatial heterogeneity. Direct study of pathogens may not be possible when predicting the risk of spread in as-yet uninfected populations, or where identification of infected individuals is unreliable. However, a direct association of host or vector gene flow with spread of the microparasite should not be assumed given the potential influence of other factors such as other host and/or vector species, environmental persistence, pathogen reproductive mode, or simply transmission via movement of non-reproducing hosts (Mazé-Guilmo et al., 2016; Tesson et al., 2016). For example, Lee et al. (2012) showed disassociation between host and virus genetic structure owing to host movement events that did not result in host gene flow (reproduction), but did result in transmission of feline immunodeficiency virus in bobcats. Indeed, correlation between host dispersal and parasite genetic structure is often weak (Mazé-Guilmo et al., 2016). Therefore, studies using host or vector data alone have limitations for inferring or predicting pathogen spread, or lack thereof, directly. However, host landscape genetic studies can provide indications of the potential risk of spread of infectious agents, and the understanding gained about host movements can inform subsequent studies of pathogen dynamics.

2.3.3 Relating spatial heterogeneity in infection risk with host spatial genetic variation

Spatial variation in pathogen prevalence or infection risk can be represented in much the same way as any landscape variable (Escobar et al., 2017), making spatial data relating to presence of an infectious agent well-suited for incorporation into host landscape genetic models. While spatial heterogeneity in pathogen prevalence could also be considered a component of the landscape that may influence spatial genetic variation in the host, typically

only adaptive loci are investigated in this context. More commonly, host neutral genetic variation is used to explain spatial patterns of infection risk or prevalence. A prominent example is a study of chronic wasting disease (CWD) in white-tailed deer. Blanchong et al. (2008) found that populations with lower CWD prevalence showed higher genetic differentiation from those that had high CWD prevalence. This genetic differentiation was found to be associated with roads and rivers, which were likely barriers to both host gene flow and CWD spread. These inferences have subsequently informed and been verified by additional landscape epidemiological research (Robinson et al., 2013).

Spatial heterogeneity in pathogen infection risk can also drive microevolutionary responses in the host (Epstein et al., 2016; Monello et al., 2017). Host species are constantly being challenged by parasitic organisms, which, if not overcome, cause disease and can have fitness consequences. This can create strong selection that acts on various genes, and geographic variation in selection at loci that are known to be associated with adaptive immune genes may reflect variation in pathogen pressure, and individual infection or disease risk (Fumagalli et al., 2011). This variation may be tested for association with environmental features such as temperature, humidity, or urbanisation (Tonteri et al., 2010), enabling insights into how future changes in climate or land use might influence overall pathogen prevalence.

2.3.4 Pathogen genetic variation to quantify pathogen transmission and spread

Using the sampled disease agent as the source of genetic data is the most direct way to infer pathogen spread across landscapes, but can be challenging to accomplish. Genetic material may be absent from, or uninformative in some infectious agents, such as prions or clonally transmissible cancers, necessitating genetic analysis of the host (e.g., Kelly et al.,

2014; Storfer et al., 2017). In addition to the aforementioned difficulties with pathogen diagnosis, pathogen nucleic acid can be difficult to isolate from samples taken from the host or vector, and would ideally be present in the blood, saliva, or other easily collected sample. Samples may also require enrichment to obtain sufficient quantities of genetic material for analysis, which can be difficult to accomplish for many pathogens, particularly viruses. However, genetic information from viruses may be particularly useful for molecular epidemiologic analyses due to their rapid mutation rate that can closely infer transmission history (Archie et al., 2009; Brunker et al., 2012). Further, viruses are prominent emerging pathogens, and have relatively small genomes, aiding whole genome-analysis. Landscape effects on viral transmission are typically studied using phylogenetic approaches (e.g., Joannon et al., 2010; Streicker et al., 2016; Fountain-Jones et al., 2017a; Young et al., 2017). To date, pathogens with larger and more slowly-mutating genomes, such as protozoans (e.g., Carrel et al., 2015; Lo et al., 2017) and fungi (e.g., Rieux et al., 2013; Brar et al., 2015), have been studied using population genetics-based methods with highly variable microsatellite and SNP loci. New methods based on next-generation sequencing technologies such as targeted enrichment techniques (e.g., Lee et al., 2017) are helping to address challenges with sequencing viruses and other pathogens, facilitating greater use of pathogen genetic data in future landscape genetic studies.

The pathogenicity of an infectious agent and the length of its period of infection are other factors that may determine its utility for landscape genetic study. Highly pathogenic agents (i.e., those that cause a greater severity of disease) are typically of utmost interest due to their potential implications for wildlife conservation, agricultural production, and human health. Among the studies identified in our literature search, agents that form chronic infections and have moderate or high pathogenic effects on their sampled host/vector organisms were more frequently investigated (see Appendix Table A2.1). However,

genetically inferring transmission histories of pathogens in host populations experiencing high rates of mortality may be problematic because hosts through which the pathogen has spread may no longer be present in the population and are unable to be sampled. Similarly, acutely infectious agents from which the host recovers after a short period of time may also evade sampling. These characteristics may be less problematic for agricultural populations where morbidity or mortality can be identified and samples collected immediately, but inevitably leave "breaks" in the inferred chain of transmission among wild populations. While complete sampling of wild populations is rarely possible in any case, obtaining adequate sample sizes is easier for apathogenic or low-pathogenicity agents that form chronic infections and may be sampled at any time post-infection (e.g., feline immunodeficiency virus; Biek et al., 2006; Lee et al., 2012; Fountain-Jones et al., 2017a). Landscape genetic study of such "model" infectious agents may be used to target specific ecological questions and provide insights into how similarly transmitted agents with higher pathogenicity might spread in the event of an outbreak.

2.4 Common methodological approaches in landscape genetics and their use in studying pathogen dynamics

There are a variety of methods available for implementing landscape genetics, some designed specifically for landscape genetics, while others have been adapted from other fields. The rapid development of landscape genetics means that new methods are regularly emerging, and it is difficult to comprehensively review all of them. However, there are some well-established methodological approaches that have either seen wide use for some time or are becoming increasingly popular at the cutting edge of the field (Balkenhol et al., 2016a).

We describe the approaches (Table 2.1) and discuss their implementation in the study of pathogen transmission and spread.

Table 2.1. Common landscape genetic approaches and their potential use in pathogen research.

Landscape genetic approach	**Their potential applications in pathogen research**	**Pathogen landscape genetic examples**
Landscape genetic simulation modelling	Predict pathogen spread in future landscape scenarios; predict spread of genes relevant to host-pathogen-vector interactions; test and validate new methods	Rees et al. (2008), Landguth et al. (2016), Leo et al. (2016)
Clustering and assignment methods	Detect barriers to pathogen spread and infer levels of barrier permeability; detect pathogen or vector environmental niche variation	Cullingham et al. (2009), Cote et al (2012), Addis et al. (2015), Brar et al. (2015)
Landscape resistance surfaces	Identify probable transmission routes or corridors; identify hosts and vectors responsible for pathogen spread; predict effects of environmental change on pathogen spread	Liang et al. (2014), Streicker et al. (2016), Lo et al. (2017), Young et al. (2017)
Graph theory and network models	Genetic inference of host contacts; identify key habitat patches/populations contributing to pathogen spread	None.
Genomic approaches	Identify associations of known candidate loci with spatial variation in pathogen exposure; infer spatial variation in pathogen exposure in different landscapes using associated loci; identify alleles determining disease susceptibility and incorporate the distribution of these into predictions of future pathogen spread	Garroway et al. (2013), Larson et al. (2014), Roffler et al. (2016), Wenzel et al. (2016)

2.4.1 Simulation modelling to test theoretical and predicted scenarios and validate methodology

In landscape genetics, simulation models are usually agent-based and spatially explicit (Landguth et al., 2016). Genetic data is modelled for individuals which have discrete spatial locations with respect to one another and with respect to environmental heterogeneity. Individuals move, behave, and reproduce according to their own attributes in response to other individuals and in response to the simulated environment, and the model simulates

changes in allele frequencies in response to these parameters. Landscape genetic simulation modelling has been used to test and validate methodological approaches (e.g. Cushman et al., 2013; Zeller et al., 2016), address theoretical questions about how and why landscape heterogeneity influences genetics (Landguth et al., 2010), and evaluate and explain empirical observations (Shirk et al., 2012). Further, simulation modelling can predict how a system might respond to certain changes, such as habitat fragmentation or future management activities.

Simulation modelling has been widely implemented in the study of pathogenic and non-pathogenic disease, beginning with medical research in the 1960s (Elveback & Varma, 1965). Frequently, epidemiological simulations are used to predict the spread of pathogens and their effect on host populations (e.g. Calonnec et al., 2008). However, the use of landscape genetic simulations in pathogen studies has been relatively limited. Landscape genetic simulations have been used to predict raccoon rabies transmission risk across a river barrier by simulating various rates of host dispersal and comparing these outputs with empirical genetic data from the host (Rees et al., 2008). The spread of particular host genes relevant to disease can also be simulated to inform management efforts. For instance, Landguth et al. (2017) used landscape genetic simulations to determine optimal planting regimes to maximise the spread of blister rust resistant genes among whitebark pine populations. Such simulations could undoubtedly be applied to vector species in particular, such as predicting the spread of pesticide resistance genes in mosquitos (Chang et al., 2016) and selecting appropriate sites for introduction of genetically modified vectors (Lavery et al., 2008). Additionally, with the need to develop further landscape genetic frameworks for the study of pathogens, simulation modelling can prove useful in testing and validating these techniques, as it has done in the broader landscape genetics field (e.g. Cushman et al., 2013; Zeller et al., 2016). For example, Leo et al. (2016) used landscape genetic simulations to

validate their multi-taxa integrated landscape genetic framework, which appears to be a promising solution to the challenge of studying pathogens with multiple hosts and/or vectors. Landscape genetic simulations may also include epidemiological parameters such as mortality or activity responses to infection, or limited infectious periods, which may otherwise confound conventional (i.e., non-simulation) landscape genetic approaches.

2.4.2 Clustering and assignment methods for quantifying connectivity and identifying transmission origin

Landscape genetic clustering and assignment methods have largely built upon classical methods from population genetics (e.g., principal components analysis, STRUCTURE, Pritchard et al., 2000) by incorporating spatial information (e.g., GENELAND, Guillot et al., 2005; sPCA, Jombart et al., 2008) and environmental heterogeneity (e.g., constrained ordination, Anderson & Willis, 2003; POPS, Jay, 2011) into estimates of population structure and providing quantitative estimates of ancestry for each individual (François & Waits, 2016). Clustering methods have been relatively popular in studying pathogens and implemented for the inference of landscape barriers affecting both host (Cullingham et al., 2009; Frantz et al., 2009; Cote et al., 2012; Addis et al., 2015) and microparasite (Rieux et al., 2011; Brar et al., 2015) spatial genetic variation. Edge detection methods, such as Monmonier's maximum difference algorithm, (Monmonier, 1973) have also been used to detect landscape barriers to transmission in pathogen studies (Joannon et al., 2010; Carrel et al., 2015). Ancestry estimates from model-based clustering algorithms can assign individuals to their populations of origin, enabling inference of landscape barrier permeability through the identification of migrants and thus estimation of the risk of pathogen spread across the barrier.

Most of the studies implementing clustering and assignment methods did not use approaches that incorporate environmental data. Instead, spatially or non-spatially explicit methods were typically used to identify genetic discontinuities and relationships with landscape barriers were inferred ad hoc, or analyses proceeded to entirely different methods that explicitly include environmental data. Associations between genetic discontinuities and landscape barriers should be considered with care due to the potential effect of intrinsic isolation mechanisms on genetic structure. For instance, studies on apple scab identified two distantly related lineages that are reproductively isolated through host specificity, but which have formed a narrow secondary contact zone in orchards where multiple host species are cultivated (Leroy et al., 2013; Lemaire et al., 2016). Secondary contact zones occurring at equivalent spatial scales to that of landscape heterogeneity may result in genetic discontinuities resembling a barrier effect and thus be misattributed as such.

Other applications of clustering methods that explicitly integrate landscape variables, such as detecting environmental niche variation (e.g., Pease et al., 2009) and ancestry-environment relationships (e.g., Jay et al., 2012), remain relatively unexplored among studies of pathogen dynamics. These applications could translate in infectious organisms to the identification of distinct ecotypes, or identify landscape features that coincide with infection foci or sources of pathogen spread. However, it must be noted that many of the genetic clustering and assignment methods presented here rely on classical population genetics models that generally do not apply to microorganisms. Consequently, the use of such methods in the study of pathogens is often limited to the inference of pathogen movement using host gene flow, with the exception of some fungal pathogens (Rieux et al., 2011; Brar et al., 2015). Some recent methods for identifying spatial population structure are free of classical population genetic assumptions, such as LOCALDIFF (Duforet-Frebourg & Blum,

2014), EEMS (Petkova et al., 2016), and MAPI (Piry et al., 2016), and can be applied to pathogens directly without these potential constraints.

2.4.3 Resistance surface modelling can identify transmission pathways and quantify spread by hosts and vectors

Resistance surfaces are commonly used in landscape genetics for modelling hypotheses concerning the influence of landscape features (from GIS landscape variables) on functional connectivity using techniques such as least-cost paths (Adriaensen et al., 2003) or circuit theory (McRae et al., 2008). These techniques produce measures of landscape or "effective" distance among populations or individuals for each hypothesis, which can be tested against observed genetic variation. The primary applications of resistance surface modelling in landscape genetics have been the identification of dispersal corridors and predicting the impacts of landscape and environmental change, such as habitat fragmentation or climate change, on connectivity. Similarly, landscape genetic resistance surfaces can identify transmission corridors or future patterns of spread (e.g., Streicker et al., 2016), and such tools have been identified previously as having great utility for pathogen landscape genetic studies (Biek & Real, 2010). However, resistance surface modelling remains infrequently applied among pathogen studies. Careful consideration is required for identifying the most relevant landscape variables to be tested, and correctly parameterising (assigning costs to) the resistance surface(s) so that these variables are represented in a biologically meaningful way. Developing landscape resistance hypotheses for transmitted agents may be more difficult as their interactions with the landscape are often indirect, mediated by the ecology of hosts and vectors. Pathogen ecological niche models offer an empirical approach for constructing resistance surfaces based on ecological factors

influencing pathogen prevalence (Escobar et al., 2017; Fountain-Jones et al., 2017c), but these also may not adequately represent host/vector movements.

Our literature search returned only one study that explicitly modelled landscape resistance based on pathogen-specific biology, testing elevation (as a proxy for temperature) as a predictor of *Plasmodium* spread, in addition to resistance surfaces that modelled human movements and mosquito vector ecology (Lo et al., 2017). However, several other studies applied resistance surfaces to hosts and vectors. Young et al. (2017) tested resistance surfaces based on waterbird niche models against genetic data from avian influenza. Two further examples of resistance surface modelling used host genetic data: Liang et al. (2014) studied landscape resistance of the snail *Oncomelania hupensis* to infer transmission of its parasite *Schistosoma japonicum*; while Rioux Paquette et al. (2014) identified likely dispersal corridors for two rabies hosts. Further, Streicker et al. (2016) used resistance surfaces to construct least-cost pathways predicting future spread of vampire bat rabies, and Barton et al. (2010) tested landscape resistance to rabies gene flow among striped skunks using landscape variables believed important for host dispersal, as well as landscape variables found to be relevant to other rabies hosts. Resistance surface modelling has thus demonstrated utility in identifying landscape drivers of functional connectivity that can shape pathogen spread and should see increasing use as frameworks for the integration of host, vector, and pathogen data continue to develop.

2.4.4 Graph theory and network models – integrating landscape genetic and epidemiological approaches

Graph theoretic approaches, which describe connections (edges) between discrete objects (nodes) (Newman, 2003), are a flexible yet powerful tool for use in landscape

genetics (Garroway et al., 2008; Dyer et al., 2010). In landscape genetics, nodes can represent individuals, populations, or habitat patches, possessing genetic parameters such as diversity measures (e.g., Dyer et al., 2010), or landscape parameters such as percentage habitat or habitat quality (e.g., Murphy et al., 2010a). Similarly, edges can represent genetic relationships between nodes such as genetic distances, gene flow, or dispersal (e.g., Decout et al., 2012), or spatial/landscape relationships such as geographic distance or landscape resistance (e.g., Dyer et al., 2010). Distinct from other landscape genetic analytical approaches, graphs allow inferences based on the overall shape, or topology, of the network, which can provide unique insights into system-wide processes, such as hierarchical population structure (Dyer & Nason, 2004).

Network topology may be used to identify populations or habitat patches that form important "stepping stones" for maintaining genetic connectivity across an entire system. Such an approach enables experimental simulation whereby nodes may be selectively removed and the overall effect on the system's topology (e.g. overall connectivity, population structure) assessed. Metrics pertaining to the importance of individual nodes to network topology can be correlated with variables such as landscape to identify important drivers of network processes. Despite their unique applications, graph theory and network approaches are relatively under-utilised in landscape genetics compared to methods specifically derived from population genetics and landscape ecology. However, among studies of infectious agents, network approaches in wildlife are becoming increasingly popular (Craft & Caillaud, 2011; Craft, 2015). Epidemiological network models are typically based on host contact networks, which are usually constructed using direct observations or indirect techniques such as mark-recapture, telemetry, or proximity loggers, and pathogens are simulated on these contact networks. Such approaches have already incorporated landscape and other environmental features. Additionally, the potential for inferring host contacts in network

models using pathogen genetic markers (see below) has been acknowledged in recent reviews (Craft, 2015; White et al., 2017; Gilbertson et al., 2018), and some studies have directly compared host contact network parameters to parasite genotypes (e.g., Bull et al., 2012). Despite this, to our knowledge, no published studies have used network models to investigate pathogen movement within a landscape genetic framework.

2.4.5 Genomic approaches to study microevolutionary responses to pathogens and landscape structure

While landscape genetics initially was used to investigate spatial genetic patterns using relatively few neutral markers, the more modern advent of landscape genomics allows the study of variation across the entire genome and effectively expands the scope of landscape genetics to include the study of functional, adaptive genetic variation. Next-generation sequencing (NGS) techniques such as restriction-site-associated DNA sequencing (RADseq) require minimal prior knowledge of the genome under study and can genotype thousands of SNPs randomly distributed across the genome. Some of these SNPs will by chance be located within or near (and thus linked to) genes or regulatory regions that are under selection. Genome-wide association studies (GWAS) can make use of this information to identify loci linked to phenotypic variation such as disease susceptibility. Genotyping of candidate loci identified using quantitative trait locus mapping and GWAS can be expanded across a large number of individuals using methods such as targeted sequence capture (Grover et al., 2012), and these data can be tested in a landscape genomic framework for associations with environmental variables.

Loci exhibiting a signature of selection can be identified using outlier tests (e.g., Excoffier et al., 2009; Luu et al., 2017), which search for loci with allelic frequencies that are

outliers relative to the majority. Such loci are considered potentially under selection and may then be tested *a posteriori* for correlations with environmental variables. Newer methods have focused on explicitly incorporating environmental variables into landscape genomic analyses, known as genetic-environment association (GEA) tests (Lotterhos & Whitlock, 2015; Rellstab et al., 2015). GEA analyses test for correlations between environmental variables and individual genotypes, which eliminates problems due to underlying population structure that must be controlled when using outlier tests. NGS approaches also generate thousands of neutral loci, which provide greater power to detect fine-scale neutral genetic structure than conventional studies based on relatively few loci (Allendorf et al., 2010). However, for studies with a particular focus on functional genetic variation, NGS approaches can also be adapted specifically for this purpose through targeted sequencing of the exome (e.g., Roffler et al., 2016) or transcriptome (e.g., de Wit & Palumbi, 2013).

While genomic technologies are becoming the norm in microbial research, approaches incorporating landscape (i.e. landscape genomics) have yet to see widespread usage. Current examples focus on established candidate loci from the host known to be relevant to immune function from prior research, such as in commercially important salmon (Tonteri et al., 2010; Larson et al., 2014). Wenzel et al. (2016) used SNPs previously identified to be associated with nematode burden in red grouse (*Lagopus lagopus scotica*) using GWAS, as well as F_{st} outlier loci, to investigate parasite-driven genetic structure across a landscape. Garroway et al. (2013) used SNPs identified from transcriptome sequencing of great tits (*Parus major*) to conduct a GEA with respect to avian malaria infection risk and contrasted this with neutral gene flow. Another targeted approach employed exon capture to genotype SNPs that were then subjected to outlier and environmental association tests, to investigate variation in pathogen exposure with respect to environment (Roffler et al., 2016). The spread of functional alleles has also been incorporated into landscape genetic simulations (Landguth et

al., 2017), enhancing predictions of future pathogen spread and its effects on host populations. This small body of research is promising for expansion of landscape genomic studies designed to couple pathogen-related functional genetic variation with landscape variables.

2.5 Emerging concepts for the landscape genetics of infectious agents

While we believe that there remains much unexplored utility in established landscape genetic methods for the study of pathogen dynamics as we have described above, we also note new frontiers with significant potential for expanding research in this area. We complete this review by discussing three particularly promising frontiers.

2.5.1 Simultaneously integrating host, vector, and landscape variables into studies of pathogen gene flow

Studies relating pathogen genetic data directly to the landscape using resistance surfaces are challenged by the mediating influence of distinct host and vector traits, as well as relative differences in the contributions of multiple host and/or vector species to microparasite gene flow. This necessitates frameworks that more holistically incorporate multiple host and vector factors into studies of pathogen gene flow, which can expand the potential insights provided by landscape genetic studies of infectious agents (Figure 2.2). Single or multiple host or vector species can be added as "landscape variables" (e.g., as resistance surfaces) in addition to physical landscape and environmental variables to test as factors shaping spatial pathogen genetic structure. Resistance surfaces for tests of microparasite gene flow can represent host/vector distributions or abundance, ideally inferred from empirically derived ecological niche or species distribution models. Optimally,

host/vector movement would be represented (Dougherty et al., 2018), using outputs from agent-based movement models informed by telemetry or mark-recapture data, or host/vector landscape genetic data representing spatial patterns of gene flow. We note that the common issue in conventional landscape genetics of spatio-temporal mismatches between landscape processes and genetic change (Anderson et al., 2010; Landguth et al., 2010) would apply even more strongly here. Researchers must simultaneously consider the potentially different spatial and temporal scales over which host and pathogen genetic changes (and potentially those of additional host/vector species) and landscape changes occur. Nonetheless, with careful study design, such a framework has significant potential to unify hosts/vectors and landscape variables under a single analytical framework for explaining and predicting pathogen transmission and spread. Importantly, it allows a flexible framework for both single and multi-host/vector systems.

Approaches integrating multiple host and vector datasets into landscape genetic studies of infectious agent gene flow have been proposed recently. Leo et al. (2016) developed a multi-taxa integrated landscape genetic framework for diseases, which simultaneously quantifies the effects of both landscape variables and interspecific co-dispersal on pathogen gene flow in multi-host-vector systems. Few studies include both host and pathogen genetic data in landscape analyses. However, Fountain-Jones et al. (2017a) tested host relatedness, in addition to host demographic and landscape variables, and found this to be an important predictor of pathogen gene flow in a single-host system of feline immunodeficiency virus in bobcats. A framework for combining multiple host and vector resistance surfaces was recently demonstrated for Chagas disease by Schwabl et al. (2017), involving subsequent validation using landscape genetic simulations. Although their approach does not explicitly include host and vector movement or gene flow, it advocates careful selection of landscape variables based on each host and vector species, informed by

previous research. A recent study tested least-cost path models of water bird movement estimated from ecological niche models, and road networks representing human movement, as potential predictors of avian influenza spread (Young et al., 2017). Other studies have investigated landscape genetic structure in multiple hosts of the same pathogens, identifying divergent dispersal patterns that could be integrated into studies of pathogen gene flow under such a framework (Vander Wal et al., 2013; Rioux Paquette et al., 2014). Approaches that consider whole ecological communities have recently been identified as necessary for advancing our understanding of pathogen dynamics (Johnson et al., 2015; Fountain-Jones et al., 2017c). Studies integrating multiple host and vector species into landscape genetic models of spread of infectious agents represent an important step towards such a paradigm.

2.5.2 Using molecular markers from infectious agents to detect cryptic landscape-host processes

The rapid mutation of microparasites relative to their hosts has potential to provide greater power to detect subtle variation in host movement patterns in response to the landscape, as well as earlier detectability of changes in host movements (such as in response to a new barrier) that are yet to be reflected in host genetic structure (Landguth et al., 2010). Additionally, movements of non-reproducing hosts are difficult to detect using host genetic markers, but instead might be inferred using markers from directly transmitted microorganisms. Such an approach has demonstrated utility of a chronic, relatively apathogenic infection of felids (feline immunodeficiency virus) for identifying demographic structure of mountain lions and recent population history (Biek et al., 2006), and has identified movement of bobcats across a highway barrier that was not detectable using host

markers (Lee et al., 2012). However, these approaches have not been broadly applied, particularly in the study of landscape effects.

The application of microparasite molecular markers to the study of host movements should be considered with some caveats in mind. As we have discussed previously, host movements and parasite transmission are not necessarily coupled (Mazé-Guilmo et al., 2016), and assumptions about how closely parasite gene flow may reflect host movements must be made with care. Also, rapid mutation and rapid dissemination of new alleles (i.e., infectious agent transmission can generally occur much more quickly than host reproduction) largely limits the use of genetic data from the microparasite to the study of very recent or ongoing landscape changes, while older processes are better studied using host genetic data. However, choice of infectious agent based on its characteristics (e.g., mutation rate, prevalence in a population, mode of transmission), may be tailored to the type and age of the host process under study. We thus do not propose that microparasite genetic data alone should be used to study host movements in their entirety, but rather that it may have specific utility as a complementary approach to host markers for providing a more complete analysis of host movement. Such insights might include contacts between specific hosts, potentially including interspecific interactions. As new infectious agents are discovered and their relationships with host movements become better understood, microparasite molecular markers will have increasing utility in landscape genetic research of host populations.

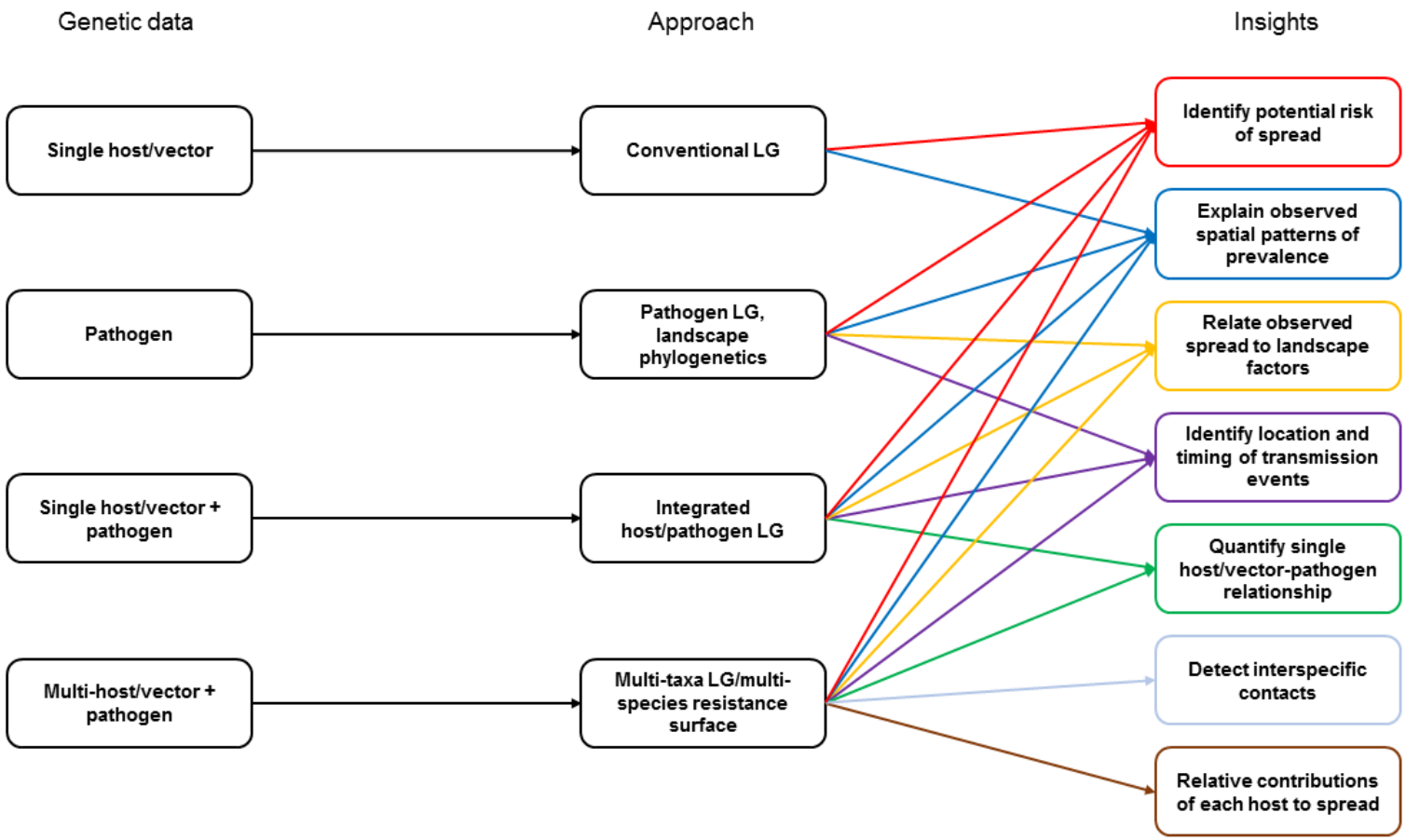

Figure 2.2. Schematic indicating the increasing insights that may be gained from using approaches that are able to integrate additional host/vector and pathogen genetic datasets, starting with genetic data from a single host or vector, through to multi-species approaches that integrate multiple host, vector, and pathogen datasets. LG = landscape genetic

2.5.3 The role of phylogenetics in understanding landscape influences on pathogen genetic variation

Phylogenetic approaches can reconstruct very recent epidemic histories, providing insights into particular transmission events and pathways that may be contextualised temporally and spatially (e.g., Corman et al., 2014; Faria et al., 2014; Carroll et al., 2015; Magee et al., 2015; Fountain-Jones et al., 2017c, 2017b). The majority of such work has been conducted on RNA viruses owing to their small, rapidly mutating genomes, requiring relatively little sequencing effort to detect contemporary phylogenetic signals. Other pathogens that evolve more slowly, such as bacteria or fungal pathogens, require the sequencing of larger portions of their genomes to capture equivalent phylogenetic signals (Biek et al., 2015). While this is becoming increasingly feasible (Kao et al., 2014), more complex computational analysis is required to make meaningful conclusions.

Several approaches may be used for relating phylogenetic information with landscape variables. Neighbour joining trees can identify clusters for quantifying population-level landscape genetic relationships (Joannon et al., 2010). The calculation of genetic distances based on maximum likelihood trees (e.g., Real et al., 2005; Carrel et al., 2012; Young et al., 2017) result in distance matrices that can be correlated with landscape resistance matrices using conventional landscape genetic approaches. Relaxed random walk phylogeographic approaches (Lemey et al., 2010) that can reconstruct pathogen dispersal have been linked to landscape predictors using a "phylogeographic GLM" method (Faria et al., 2013; Jacquot et al., 2017). The phylogeographic GLM approach has enabled a better understanding of how landscape and hosts can constrain pathogen spread. For example, using the phylogeographic GLM approach on viral genomic data, roads and rivers, coupled with dog distribution, were found to impact rabies spread in Tanzania (Brunker et al., 2018). However, this approach is

limited to discrete sampling locations and is computationally intensive (Dellicour et al., 2016b). A recent framework by Dellicour et al. (2016b) modifies the phylogeographic GLM approach to use resistance surfaces to efficiently quantify landscape resistance along transmission pathways inferred by continuous phylogeographic analyses. These landscape resistances are then correlated with temporal estimates of transmission along these routes to estimate how the landscape has shaped rates and directions of pathogen spread. Such approaches are yet to be broadly applied, but appear to be important developments that should see increasing application in the future.

2.6 Conclusion

Overall, landscape genetics has been relatively under-utilised in disease ecology research. We believe this is partly due to a lack of cross-disciplinary awareness between the two fields, but also a lack of a clear landscape genetic framework specifically designed for tackling pathogen systems, which are often complex and do not facilitate easy translation of existing landscape genetic tools. However, we note there has been a recent effort to develop new frameworks for such research, expanding the utility of the landscape genetic toolset. These tools will increase our capacity to study complex multi-host and host-vector systems, improving the integration of multiple genetic datasets and accounting for interspecific interactions. Improved understanding of host-parasite associations will facilitate the use of microparasite genetic markers to provide insights into host processes that may be difficult to detect using conventional host landscape genetics. Identification of idealised systems that are designed to target specific ecological questions will also facilitate progress in this field. Recent methods that enable the incorporation of quantitative landscape data into spatio-

temporal phylogenetic reconstructions of recent transmission events, coupled with advances in high-throughput sequencing, hold great promise for studying how the landscape shapes transmission processes. We believe that these recent developments represent a renewed interest in advancing landscape genetic research in pathogen systems, which we expect will translate to continued growth of research in this area.

2.7 Appendix

Table A2.1. Results retained from literature search for empirical papers using landscape genetics to address questions related to pathogens

Publication	Sampled study species	Pathogen/parasite	Pathogen/parasite type	Source of genetic data	Pathogenicity	Chronic/Acute infection
Addis, B.R., et al. (2015). Population genetic structure and disease in montane boreal toads: more heterozygous individuals are more likely to be infected with amphibian chytrid. Conservation Genetics, 16, 833-844.	Boreal toads, *Bufo boreas*	*Batrachochytrium dendrobatidis*	Fungus	Host	High	Chronic
Blanchong, J.A., et al. (2008). Landscape genetics and the spatial distribution of chronic wasting disease. Biology Letters, 4, 130-133.	White-tailed deer, *Odocoileus virginianus*	Chronic wasting disease	Prion	Host	Moderate	Chronic
Brar, S., et al. (2015). Colonization history, host distribution, anthropogenic influence and landscape features shape populations of white pine blister rust, an invasive alien tree pathogen. Plos One, 10, e0127916	White pine, *Pinus strobus*	White pine blister rust, *Cronartium ribicola*	Fungus	Pathogen	High	Chronic
Campbell, L.P., et al. (2017). Landscape genetics of *Aedes mcintoshi* (Diptera: Culicidae), an important vector of Rift Valley fever virus in Northeastern Kenya. Journal of Medical Entomology, 54, 1258-1265.	*Aedes mcintoshi*	Rift Valley fever virus	Virus	Vector	Low	Chronic
Carrel, M., et al. (2015). The geography of malaria genetics in the Democratic Republic of Congo:	Humans, *Homo sapiens*	Malaria, *Plasmodium falciparum*	Protozoan	Pathogen	High	Acute

A complex and fragmented landscape. Social Science & Medicine, 133, 233-241.						
Carrel, M.A., et al. (2012). Population-environment drivers of H5N1 avian influenza molecular change in Vietnam. Health & Place, 18, 1122-1131.	Humans, *Homo sapiens*	H5N1 Avian influenza	Virus	Pathogen	Moderate-High	Chronic
Chang, X., et al. (2016). Landscape genetic structure and evolutionary genetics of insecticide resistance gene mutations in Anopheles sinensis. Parasites & Vectors, 9, 228.	Mosquito, *Anopheles sinensis*	Malaria, *Plasmodium* spp.	Protozoan	Vector	Low	Acute
Cote, H., et al. (2012). Genetic structure and rabies spread potential in raccoons: the role of landscape barriers and sex-biased dispersal. Evolutionary Applications, 5, 393-404.	Raccoon, *Procyon lotor*	Rabies	Virus	Host	High	Chronic
Criscione, C.D., et al. (2010). Landscape genetics reveals focal transmission of a human macroparasite. Plos Neglected Tropical Diseases, 4, e665.	Humans, *Homo sapiens*	Roundworm, *Ascaris lumbricoides*	Macroparasite	Pathogen	Moderate	Chronic
Cullingham, C.I., et al. (2009). Differential permeability of rivers to raccoon gene flow corresponds to rabies incidence in Ontario, Canada. Molecular Ecology, 18, 43-53.	Raccoon, *Procyon lotor*	Rabies	Virus	Host	High	Chronic/Subacute
DeYoung, R.W., et al. (2009). Landscape-genetic analysis of population structure in the Texas Gray Fox Oral Rabies Vaccination	Gray fox, *Urocyon cinereoargenteus*	Rabies	Virus	Host	High	Chronic

zone. Journal of Wildlife Management, 73, 1292-1299.						
Egizi, A., et al. (2016). The hitchhiker's guide to becoming invasive: exotic mosquitoes spread across a US state by human transport not autonomous flight. Molecular Ecology, 25, 3033-3047.	Mosquito, *Aedes japonicus*	various	Unspecified	Vector		Chronic
Frantz, A.C., et al. (2009). Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? Journal of Applied Ecology, 46, 493-505.	European wild boar, *Sus scrofa*	Swine fever	Virus	Host	High	
Garroway, C.J., et al. (2013). Fine-scale genetic structure in a wild bird population: the role of limited dispersal and environmentally based selection as causal factors. Evolution, 67, 3488-3500.	Great tit, *Parus major*	Avian malaria, *Plasmodium circumflexum* and *P. relictum*	Protozoan	Host	Low	Acute
Guivier, E., et al. (2011). Landscape genetics highlights the role of bank vole metapopulation dynamics in the epidemiology of Puumala hantavirus. Molecular Ecology, 20, 3569-3583.	Bank vole, *Myodes glareolus*	Puumala hantavirus	Virus	Host	Low-Moderate	Chronic
Joannon, B., et al. (2010). Barriers to gene flow between emerging populations of watermelon mosaic virus in Southeastern France. Phytopathology, 100, 1373-1379.	Cucurbit crops	Watermelon mosaic virus	Virus	Pathogen	Moderate	Chronic
Kelly, A.C., et al. (2014). Genetic assessment of environmental features that influence deer dispersal: implications for prion-	White-tailed deer, *Odocoileus virginianus*	Chronic wasting disease	Prion	Host	Moderate	Chronic

infected populations. Population Ecology, 56, 327-340.						
Krtinic, B., et al. (2014). Microhabitat and spatial variation at HK isozyme loci in *Culex pipiens*: testing isolation by distance and isolation by ecology model. Bulletin of Insectology, 67, 237-246.	Mosquito, *Culex pipiens*	various	Unspecified	Vector		Chronic
Krtinic, B., et al. (2016). Integrative approach revealed contrasting pattern of spatial structuring within urban and rural biotypes of *Culex pipiens*. Journal of Applied Entomology, 140, 757-774.	Mosquito, *Culex pipiens*	various	Unspecified	Vector		
Landguth, E.L., et al. (2017). Using landscape genetics simulations for planting blister rust resistant whitebark pine in the US Northern Rocky Mountains. Frontiers in Genetics, 8, 9.	Whitebark pine, *Pinus albicaulis*	Blister rust, *Cronartium ribicola*	Fungus	Host	Moderate	
Lang, K.R. and J.A. Blanchong (2012). Population genetic structure of white-tailed deer: understanding risk of chronic wasting disease spread. Journal of Wildlife Management, 76, 832-840.	White-tailed deer, *Odocoileus virginianus*	Chronic wasting disease	Prion	Host	Moderate	Chronic
Larson, W.A., et al. (2014). Signals of heterogeneous selection at an MHC locus in geographically proximate ecotypes of sockeye salmon. Molecular Ecology, 23, 5448-5461.	Sockeye salmon, *Oncorhynchus nerka*	various	Unspecified	Host		Chronic
Li, S.Z., et al. (2009). Landscape genetics: the correlation of spatial and genetic distances of *Oncomelania hupensis*, the	Snail, *Oncomelania hupensis*	*Schistosoma japonicum*	Macroparasite	Host	Moderate	

intermediate host snail of *Schistosoma japonicum* in mainland China. Geospatial Health, 3, 221-231.						
Liang, L., et al. (2014). Wetlands explain most in the genetic divergence pattern of *Oncomelania hupensis*. Infection Genetics and Evolution, 27, 436-444.	Snail, *Oncomelania hupensis*	*Schistosoma japonicum*	Macroparasite	Host	Moderate	Chronic
Lo, E., et al. (2017). Frequent spread of *Plasmodium vivax* malaria maintains high genetic diversity at the Myanmar-China border, without distance and landscape barriers. Journal of Infectious Diseases, 216, 1254-1263.	Humans, *Homo sapiens*	Malaria, *Plasmodium vivax*	Protozoan	Pathogen	High	Chronic
Lo, E., et al. (2017). Transmission dynamics of co-endemic *Plasmodium vivax* and *P-falciparum* in Ethiopia and prevalence of antimalarial resistant genotypes. Plos Neglected Tropical Diseases, 11, e0005806.	Humans, *Homo sapiens*	Malaria, *Plasmodium vivax* and *P. falciparum*	Protozoan	Pathogen	High	Chronic
Medley, K.A., et al. (2015). Human-aided and natural dispersal drive gene flow across the range of an invasive mosquito. Molecular Ecology, 24, 284-295.	Asian tiger mosquito, *Aedes albopictus*	various	Unspecified	Vector		Chronic
Montarry, J., et al. (2009). Spatio-temporal distribution of *Erysiphe necator* genetic groups and their relationship with disease levels in vineyards. European Journal of Plant Pathology, 123, 61-70.	Grape vines, *Vitis vinifera*	Grapevine powdery mildew, *Erysiphe necator*	Fungus	Pathogen	Moderate-High	

Mullins, J., et al. (2014). The influence of habitat structure on genetic differentiation in red fox populations in north-eastern Poland. Acta Theriologica, 59, 367-376.	Red fox, *Vulpes vulpes*	various	Unspecified	Host		Chronic
Real, L.A., et al. (2005). Unifying the spatial population dynamics and molecular evolution of epidemic rabies virus. Proceedings of the National Academy of Sciences of the United States of America, 102, 12107-12111.	Red fox, *Vulpes vulpes*	Rabies	Virus	Pathogen	High	
Rees, E.E., et al. (2009). Landscape modelling spatial bottlenecks: implications for raccoon rabies disease spread. Biology Letters, 5, 387-390.	Raccoon, *Procyon lotor*	Rabies	Virus	Host	High	Chronic
Richardson, J.L., et al. (2017). Using fine-scale spatial genetics of Norway rats to improve control efforts and reduce leptospirosis risk in urban slum environments. Evolutionary Applications, 10, 323-337.	Norway rats, *Ratus norvegicus*	Leptospirosis, *Leptospira*	Bacterium	Vector	Low-Moderate	Chronic
Rieux, A., et al. (2011). Inferences on pathogenic fungus population structures from microsatellite data: new insights from spatial genetics approaches. Molecular Ecology, 20, 1661-1674.	Plantain banana, *Musa × paradisiaca*	*Mycosphaerella fijiensis*	Fungus	Pathogen	Moderate	Chronic
Rieux, A., et al. (2013). Recent range expansion and agricultural landscape heterogeneity have only minimal effect on the spatial genetic structure of the plant	Plantain banana, *Musa × paradisiaca*	*Mycosphaerella fijiensis*	Fungus	Pathogen	Moderate	Chronic

pathogenic fungus *Mycosphaerella fijiensis*. Heredity, 110, 29-38.						
Rinaldi, P.A., et al. (2017). Genetic variability of *Phyllosticta ampelicida*, the agent of black rot disease of grapevine. Phytopathology, 107, 1406-1416.	Grape vines, *Vitis* spp.	Black rot disease, *Phyllosticta ampelicida*	Fungus	Pathogen	High	Chronic
Rioux Paquette, S., et al. (2014). Modelling the dispersal of the two main hosts of the raccoon rabies variant in heterogeneous environments with landscape genetics. Evolutionary Applications, 7, 734-749.	Raccoon, *Procyon lotor* and striped skunk, *Mephitis mephitis*	Rabies	Virus	Host	High	Chronic
Robinson, S.J., et al. (2012). The walk is never random: subtle landscape effects shape gene flow in a continuous white-tailed deer population in the Midwestern United States. Molecular Ecology, 21, 4190-4205.	White-tailed deer, *Odocoileus virginianus*	Chronic wasting disease	Prion	Host	Moderate	Chronic
Roffler, G.H., et al. (2016). SNP discovery in candidate adaptive genes using exon capture in a free-ranging alpine ungulate. Molecular Ecology Resources, 16, 1147-1164.	Dall's sheep, *Ovis dalli dalli*	various	Unspecified	Host		Chronic
Rosales-Castillo, J., et al. (2011). Genetic diversity and population structure of *Escherichia coli* from neighboring small-scale dairy farms. Journal of Microbiology, 49, 693-702.	Domestic cows, *Bos taurus*	*Eshcerichia coli*	Bacterium	Pathogen	Low	
Russo, I.R.M., et al. (2016). Landscape determinants of fine-scale genetic structure of a small rodent in a heterogeneous	Natal multimammate mouse, *Mastomys natalensis*	Lassa virus	Virus	Vector	Low	Chronic

landscape (Hluhluwe-iMfolozi Park, South Africa). Scientific Reports, 6, 29168.						
Schmidt, T.L., et al. (2017). Genome-wide SNPs reveal the drivers of gene flow in an urban population of the Asian Tiger Mosquito, *Aedes albopictus*. Plos Neglected Tropical Diseases, 11, e0006009.	Asian tiger mosquito, *Aedes albopictus*	various (at least dengue fever and chikungunya)	Unspecified	Vector	Low	
Sprehn, C.G., et al. (2015). Landscape genetics of *Schistocephalus solidus* parasites in Threespine Stickleback (*Gasterosteus aculeatus*) from Alaska. Plos One, 4, e0122307.	Threespine stickleback, *Gasterosteus aculeatus*	Tapeworm, *Schistocephalus solidus*	Macroparasite	Pathogen	Low-Moderate	Chronic
Storfer, A., et al. (2017). Landscape genetics of the Tasmanian devil: implications for spread of an infectious cancer. Conservation Genetics, 18, 1287-1297.	Tasmanian devil, *Sarcophilus harrisii*	Devil facial tumour disease	Transmissible cancer	Host	High	Chronic
Talbot, B., et al. (2012). Lack of genetic structure and female-specific effect of dispersal barriers in a rabies vector, the striped skunk (*Mephitis mephitis*). Plos One, 7, e49736.	Striped skunk, *Mephitis mephitis*	Rabies	Virus	Host	High	Chronic
Talbot, B., et al. (2017). Comparative analysis of landscape effects on spatial genetic structure of the big brown bat and one of its cimicid ectoparasites. Ecology and Evolution, 7, 8210-8219.	Big brown bat, *Eptesicus fuscus*	*Cimex adjunctu*	Macroparasite	Host & Pathogen	Low	Acute
Tonteri, A., et al. (2010). Beyond MHC: signals of elevated selection pressure on Atlantic salmon (*Salmo*	Atlantic salmon, *Salmo salar*	*Gyrodactylus salaris*	Macroparasite	Host	Moderate-High	Chronic/Subacute

salar) immune-relevant loci. Molecular Ecology, 19, 1273-1282.						
Vander Wal, E., et al. (2012). Influence of landscape and social interactions on transmission of disease in a social cervid. Molecular Ecology, 21, 1271-1282.	Elk, *Cervus canadensis*	Bovine tuberculosis, *Mycobacterium bovis*	Bacterium	Host	Moderate-High	Chronic
Vander Wal, E., et al. (2013). Juxtaposition between host population structures: implications for disease transmission in a sympatric cervid community. Evolutionary Applications, 6, 1001-1011.	White-tailed deer, *Odocoileus virginianus* and elk, *Cervus canadensis*	Bovine tuberculosis, *Mycobacterium bovis*	Bacterium	Host	Moderate-High	Chronic
Wenzel, M.A., et al. (2016). The role of parasite-driven selection in shaping landscape genomic structure in red grouse (*Lagopus lagopus scotica*). Molecular Ecology, 25, 324-341.	Red grouse, *Lagopus lagopus scotica*	Nematode *Trichostrongylus tenuis*	Macroparasite	Host	Moderate	Chronic
Wilder, A.P., et al. (2015). Population genetic structure of a common host predicts the spread of white-nose syndrome, an emerging infectious disease in bats. Molecular Ecology, 24, 5495-5506.	Little brown myotis, *Myotis lucifugus*	White-nose syndrome, *Pseudogymnoascus destructans*	Fungus	Host	High	Chronic
Young, S.G., et al. (2017). How's the Flu Getting Through? Landscape genetics suggests both humans and birds spread H5N1 in Egypt. Infection Genetics and Evolution, 49, 293-299.	Poultry	H5N1 Avian influenza	Virus	Pathogen	High	Acute

Chapter 3

Urbanization reduces genetic connectivity in bobcats (*Lynx rufus*) at both intra- and inter-population spatial scales

3.1 Abstract

Urbanisation is a major factor driving habitat fragmentation and connectivity loss in wildlife. However, the impacts of urbanisation on connectivity may vary among species and even populations due to specific landscape characteristics, and because the importance of different landscape factors may vary among different spatial scales. Therefore, our ability to detect these processes may depend on the spatial scale under study. Bobcats (*Lynx rufus*) are relatively sensitive to urbanisation and considered important indicators of connectivity in highly urban coastal southern California. We genotyped 278 bobcats at 13,520 SNP loci to conduct a replicated landscape resistance analysis in five genetically distinct populations, which we identified from this region. Populations varied in spatial area and landscape composition. We tested urban and natural factors potentially influencing individual connectivity in each population separately as well as region-wide. Overall, landscape genomic effects were most frequently detected at the region-wide spatial scale, with urban land cover (measured as impervious surface) having negative effects and topographic roughness having positive effects on gene flow. The negative effect of urban land cover on connectivity was also evident when populations were analysed separately despite varying substantially in spatial area and the amount of urban development, confirming a pervasive impact of urbanisation largely independent of spatial scale. The effect of urban development was strongest in one population where stream habitat had been lost to development, suggesting that riparian corridors may help mitigate reduced connectivity in urbanising areas.

Our results demonstrate the importance of replicating landscape genetic analyses across populations and considering how landscape genetic effects may vary with spatial scale and local landscape structure.

3.2 Introduction

Urban development causes habitat degradation and fragmentation (Ramalho & Hobbs, 2012). Habitat fragmentation exposes organisms to edge effects potentially leading to increased anthropogenic disturbance or changes in interspecific interactions (Murcia, 1995; Fahrig, 2003). Habitat fragmentation also isolates populations and reduces functional connectivity, defined as the degree to which the landscape facilitates or impedes movement among patches (Taylor et al., 1993). Isolated populations are susceptible to inbreeding depression and genetic drift that reduce overall fitness and adaptive potential in the face of current and future threats, such as climate change and novel pathogens (Keyghobadi, 2007; Hoffmann et al., 2017), and suffer reduced potential for demographic rescue (Brown & Kodric-Brown, 1977). However, the consequences of fragmentation may vary among populations and species due to variation in factors such as patch size, the distribution and intensity of urban development, and underlying landscape characteristics such as topography and vegetation. In addition, the observed impacts of habitat fragmentation on connectivity may vary according to the spatial scale under study, with different landscape drivers of connectivity potentially varying in importance among different (e.g., locally within populations vs. regionally among populations) spatial scales (Cushman & Landguth, 2010a). Thus, the consideration of multiple spatial scales is critical to a detailed and comprehensive understanding of urban impacts on connectivity.

The impacts of urban development can be particularly acute for mammalian carnivores (Randa & Yunger, 2006; Fuller et al., 2010; Ordeñana et al., 2010). Many carnivore species are territorial, exist at low population densities, and require large, connected areas of habitat to support viable populations. Consequently, habitat fragmentation and its genetic and demographic effects are frequently implicated in carnivore declines (e.g., Beier, 1993; Roelke et al., 1993; Lu et al., 2001; Dixon et al., 2007). Given their sensitivity to fragmentation, carnivores are excellent indicator species of functional landscape connectivity (Noss et al., 1996). Moreover, their large home ranges frequently make them useful umbrellas for conserving broader ecological communities (Noss et al., 1996). These factors, in addition to the potentially important role that carnivores can play in ecosystems (Estes et al., 2011), particularly in constrained urban landscapes (Crooks & Soulé, 1999; Crooks et al., 2010), emphasise the value in understanding how the landscape influences connectivity for carnivores.

Coastal southern California is one of the most urbanised landscapes in North America, having experienced rapid human population growth and expansion of developed areas over the past several decades (U.S. Census Bureau, 2010), with a human population of over 13.3 million in the Los Angeles metropolitan area alone (U.S. Census Bureau, 2016). This region is also renowned as a hotspot for biodiversity and endemism, with habitat fragmentation and loss leading to high concentrations of threatened species in remaining natural areas (Dobson et al., 1997; Myers et al., 2000). Despite extensive urban development, coastal southern California retains relatively intact communities of carnivore species that vary in their requirements for patch size and habitat quality and in their overall sensitivity to urbanisation (Crooks, 2002; Ordeñana et al., 2010). In particular, bobcats (*Lynx rufus*) are the third largest

carnivore in coastal southern California and are regarded as important indicators of functional landscape connectivity in this region (Crooks, 2002).

Telemetry studies show that bobcats in coastal southern California rely on natural areas that consist predominantly of coastal sage scrub and chaparral vegetation (Riley et al., 2003). However, bobcats are also habitat generalists and may persist near to and even within anthropogenically altered and populated areas (Riley et al., 2003; Lyren et al., 2008b). Although major roads and urban development are barriers to functional connectivity in this region, telemetry and genetic studies (including pathogen genetics) have indicated occasional crossing of major roads by bobcats, mostly facilitated by culverts or underpasses (Riley et al., 2006; Lyren et al., 2008b; Lee et al., 2012; Poessel et al., 2014; Serieys et al., 2015; Fountain-Jones et al., 2017a). Nonetheless, several independent microsatellite studies have broadly characterised a collection of genetically distinct bobcat populations, which are confined to discrete habitat patches of varying size separated by major roads and areas of concentrated urban development (Riley et al., 2006; Lee et al., 2012; Ruell et al., 2012; Serieys et al., 2015). While these barriers to gene flow between populations are relatively well-understood, the drivers of connectivity within populations in habitat patches separated by major roads are generally unknown.

Understanding how landscape factors within habitat patches are influencing connectivity within individual populations can provide different insights compared to the study of how the landscape between patches is influencing connectivity among populations. Factors that influence connectivity within populations are likely to differ from one population to another depending on the local landscape characteristics, such as topography or the degree of urbanisation, that can in turn determine the distribution and abundance of potentially important features for connectivity such as vegetation or riparian zones (Short Bull et al.,

2011). Furthermore, by narrowing our focus to populations and patches contained within the bounds determined by hard anthropogenic barriers such as highways (thereby excluding the strong genetic signals of these barriers from analysis), we can assess finer scales of both spatial genetic variation and landscape heterogeneity. Investigating this finer-scale variation can enable us to better detect which factors might be important for maintaining the degree of connectivity that necessarily must exist within a population, and to assess the potential impacts of landscape change on that connectivity (Cushman & Landguth, 2010a). Thus far, landscape factors influencing genetic variation within bobcat populations have not been characterised, and it remains unclear which specific natural features, if any, are important in maintaining bobcat genetic connectivity.

Next-generation sequencing technologies have greatly enhanced our ability to accurately estimate neutral genomic variation compared to microsatellites (Santure et al., 2010; Helyar et al., 2011; Fischer et al., 2017). Coupling genomic data with rigorous landscape genetic approaches that incorporate the individual (as opposed to the population) as the statistical unit provides considerable power for identifying genetic variation at fine spatial scales to quantify functional connectivity (Manel et al., 2003; Holderegger & Wagner, 2008; Cushman & Landguth, 2010b). Furthermore, replicating landscape genetic studies across multiple areas and investigating landscape genetic relationships at multiple spatial scales are important in assessing the generality of landscape effects on functional connectivity (Anderson et al., 2010; Short Bull et al., 2011; Robertson et al., 2018). The spatial structure of bobcat populations in coastal southern California is well-suited to a design that tests factors affecting functional connectivity at multiple spatial scales across several comparable, but distinct landscapes, with varying habitat structure and degrees of urbanisation (Ruell et al., 2012).

Here, we used a landscape genomic approach to identify landscape factors either promoting or constraining bobcat dispersal in coastal southern California, and to understand how these factors vary among populations from habitat patches with different landscape characteristics, including varying degrees of urbanisation. Using next-generation sequencing to genotype bobcats at 13,520 SNP loci, we identify five genetically distinct populations and define for each an area with which each population is associated. Using our high-resolution SNP data to make precise measurements of genetic relatedness among individuals, we implement an individual-based landscape genomic approach testing support for landscape resistance variables representing possible effects of different landscape factors on bobcat connectivity, which we replicate among populations. We predict that factors influencing connectivity will vary among populations, and that this variation will depend on the characteristics of each patch, particularly patch size and the degree of urbanisation. We then conduct an additional landscape resistance analysis including all individuals across the region to test the prediction that factors affecting connectivity among populations will differ from those affecting connectivity within populations.

3.3 Methods

3.3.1 Sample collection

We utilised 293 bobcat blood and tissue samples collected in three study areas in southern California: northwest of Los Angeles, southeast of Los Angeles, and north and east of San Diego (Figure 3.1). San Diego samples (n = 45) were collected between 2007–2012 according to Jennings and Lewison (2013). Northwest Los Angeles samples (n = 136) were collected between 1997–2011 according to Riley *et al.* (2006) and Serieys *et al*. (2015).

Southeast Los Angeles samples (n = 112) were collected between 2002–2010 according to Lyren et al. (2006, 2008b, 2008a). All animals were sampled from a combination of live trapping and opportunistically collected carcasses of predominantly roadkill. Live animals were captured, handled, and released using protocols approved by cooperating agencies and relevant animal ethics committees (see original publications – cited above – for detailed information).

3.3.2 Laboratory procedure

We extracted genomic DNA using the Qiagen DNeasy Blood & Tissue Kit (Qiagen Inc., Valencia, CA, USA), eluting DNA in buffer EB. Agencourt Ampure XP SPRI beads (Beckman Coulter Inc., Brea, CA, USA) were used to concentrate some low-yield DNA extractions. We prepared double-digest restriction-site-associated DNA (ddRAD) libraries according to Peterson et al. (2012), using *Nla*III and *Eco*RI-HF restriction enzymes, on individual samples normalised on a within-library basis to at least 200 ng DNA in 25 µL. A fragment size of 300–380 bp (excluding 75 bp adapters) was selected using a Blue Pippin size selection system (Sage Science Inc., Beverly, MA, USA) with a 100–600 bp 2% agarose gel cartridge containing internal standards, with fragment size verified using an Agilent Tapestation 2200. Forty-eight uniquely barcoded P1 adapters enabled subsequent identification of pooled individuals, with biotinylated P2 adapters enabling streptavidin Dynabead (Invitrogen, Carlsbad, CA, USA) purification to maximise efficiency of the final PCR amplification step. PCR was conducted over 12 cycles across five reactions per pool of individuals, using uniquely indexed primers to enable identification of each individual according to the pool of which it was a part, creating a two-tiered barcode or index system.

This per-pool indexing allowed individual P1 barcodes to be used more than once among multiple pools and combined within the same library.

An initial trial library of 16 individuals was prepared, followed by three libraries of 80 individuals, and one library of 82 individuals, multiplexed to obtain a minimum average sequencing coverage of 12x per locus, per individual. To enable estimation of sequencing error rates for optimisation of locus assembly parameters (Mastretta-Yanes et al., 2015), detailed below, each of the prepared libraries contained five within-library replicates and five replicates shared with other libraries, except for the 16-sample library, which contained two individuals shared with other libraries/lanes. Collectively, a total of 12 individuals were replicated between libraries and 20 were replicated within libraries, among 306 unique individuals (including re-runs of 13 individuals due to initially low sequencing coverage). Sequencing was conducted at the University of Oregon Genomics & Cell Characterization Core Facility for 100 bp, single-end reads, firstly on an Illumina Hi-Seq 2500 for the 16-individual and one 80-individual library, and on an Illumina Hi-Seq 4000 for the subsequent 80- and 82-individual libraries. Each library was sequenced on a separate lane.

3.3.3 Bioinformatics and data filtering

Raw sequence files were initially checked for quality using FastQC (Andrews, 2010). Stacks v1.42 (Catchen et al., 2013) was then used to assemble reads into loci and identify single-nucleotide polymorphisms (SNPs). Per-individual demultiplexing of sequencing reads, Phred score quality filtering, and trimming of adapters was performed using the Stacks program *process_radtags.* The Stacks pipeline for non-reference-aligned data, *denovo_map.pl*, was used to build loci and identify SNPs from 'stacks' of sequence reads,

populate a catalogue containing sets of consensus loci, and match individuals against the catalogue to call alleles at each locus for each sample.

Four user-specified parameters have been shown to influence rates of error introduced during *denovo_map.pl*, with optimal settings being specific to each dataset (Mastretta-Yanes et al., 2015; Paris et al., 2017): the minimum number of identical, raw reads required to create a stack (-*m*); the maximum number of mismatches allowed between loci when processing a single individual (-*M*); the maximum number of mismatches allowed between loci when building the catalogue (-*n*); and the maximum number of stacks at a single *de novo* locus (-*max_locus_stacks*). Using between- and within-library replicates, we conducted 11 trials of *denovo_map.pl*, varying a single parameter at a time, and calculated locus error (proportion of loci genotyped in only one of a pair of replicate individuals), allele error (proportion of allele mismatches among replicate pairs per locus), and SNP error (proportion of SNP mismatches among replicate pairs) for each trial according to Mastretta-Yanes et al. (2015).

denovo_map.pl was run using the full dataset with parameter settings chosen to minimise error and maximise the number of SNP loci (-*m* = 3, -*M* = 2, -*n* = 4, -*max_locus_stacks* = 3). We generated a SNP matrix containing allele calls for each individual using the Stacks program *populations* with minimal filtering, except to retain only loci that were present in >20% of individuals in each population (defined in this step as northwest Los Angeles, southeast Los Angeles, and San Diego populations, as above), and only a single, randomly chosen SNP per locus. We used Plink v1.07 (Purcell et al., 2007) for further filtering of the SNP matrix. Loci missing from >35% of individuals were removed, followed by individuals missing >50% of loci, and loci with a minor allele frequency <0.01. A disproportionate number of SNPs were located at read positions 94 and 95, indicating increased sequencing error at these positions; these SNPs were also removed.

Principal components analysis revealed a batch effect pertaining to our 80-sample library sequenced on the Illumina Hi-Seq 2500, which produced lower average coverage than our other libraries. This was corrected using R (R Development Core Team, 2013) by comparing rates of missing data for each locus per library and removing loci for which the missing data rate in any library was above or below 1.5 * the interquartile range (as calculated from the missing data rates for a given locus across all libraries). We used PCAdapt (Luu et al., 2017) to identify outlier loci potentially under selection, with Q-value false discovery rate correction ($\alpha_Q = 0.10$). Following filtering, *populations* in Stacks was re-run using SNP-specific whitelists to produce final matrices containing putatively neutral SNPs that passed filtering for use in subsequent population and landscape genetic analyses.

3.3.4 Population genomic structure

Patterns of neutral genomic variation were initially visualised using principal components analysis in R. To resolve further population structure, *fastStructure* (Raj et al., 2014) was run using simple priors and five-fold cross validation to estimate the optimal number of genetic clusters (K). The *fastStructure* script *chooseK.py* was used to select a range of potentially optimal K values based on maximising marginal likelihood (tends to underestimate K; Raj et al., 2014) and choosing the minimum number of populations that have a cumulative ancestry contribution of at least 99.99% (tends to overestimate K; Raj et al., 2014). These K estimates were subsequently verified by examining cross validation errors and comparing with our *a priori* knowledge of landscape barriers (i.e., major roads previously shown to restrict gene flow). Once population genetic clusters were identified, individual ancestry assignment and calculation of admixture proportions were performed

using *fastStructure* with the above calculated optimal K and logistic priors. Assignment plots of *fastStructure* results were constructed using *ggplot2* (Wickham & Chang, 2008).

Population genetic statistics were calculated for each of the genetically defined populations following removal of first-generation migrants, which we defined as having >50% *fastStructure* assignment probability to a population other than that in which the individual was sampled. Observed and expected heterozygosity and *F*-statistics were calculated using *diveRsity* (Keenan et al., 2013). Allelic richness (A_r) was calculated using HP-RARE (Kalinowski, 2005) using the rarefaction procedure to account for differences in sample size among populations. The rarefaction sample size was set according to the population with the smallest sample size and was calculated as twice the number of individuals in this population genotyped for a given locus, averaged across all loci (we doubled the number of individuals as there were two alleles per locus). Effective population size (N_e) was calculated using the linkage disequilibrium method in NeEstimator (Do et al., 2014) while controlling for the number of chromosomes to correct for downward bias in N_e estimates resulting from linkage within chromosomes (Waples et al., 2016). Pairwise genetic distances among populations (F_{ST}) were calculated using *diveRsity* with 95% confidence intervals calculated using 1000 bootstraps carried out over individuals within populations. Pairwise matrices of individual genetic distance were calculated as the inverse of the proportion of shared alleles (D_{ps}) using *adegenet*, both for each population individually as well as the entire region-wide dataset.

3.3.5 Landscape genomics

We used a landscape resistance framework to test hypotheses regarding the influence of landscape features on gene flow (Table 3.1). To account for potential differences among populations in the drivers of gene flow, these hypotheses were tested within each population separately, but were then also tested among individuals across the entire region. ArcGIS 10.3 (ESRI) was used for all manipulation and analysis of spatial data except where specified otherwise.

We defined the landscape area associated with each genetic population according to the following methodology. Where a freeway or major highway (i.e., listed on the California State Highway network; Caltrans, 2017) passed near sampled individuals, and there were no individuals from the same genetic population sampled on the opposite side of the highway (with the exception of first generation migrants), this was used to define the spatial edge for a given population. However, where the edge of functionally impervious urban development formed a continuous barrier between sampled individuals and the nearest major highway, this was instead used to define the population edge. The urban edge was simplified and defined using a method adapted from Ruell *et al.* (2012), by calculating for each 30 m by 30 m raster cell whether a majority (>50%) of surrounding cells within a 1 km radius contained greater than 50% impervious surface (defined as medium or high density urban landcover according to the National Landcover Database). In one population where there was no clear geographic barrier with which to define one edge, we defined the edge using a minimum bounding rectangle around all sample locations with a 20 km buffer. We believe this 20 km buffer is sufficient to capture any landscape that might be utilised for dispersal among our sample locations, but we acknowledge that quantification of landscape characteristics within areas defined in this way is imprecise. Previous research indicates that despite occasional short

incursions, bobcats rarely make substantial movements through heavily urbanised areas (Riley et al., 2003; Lyren et al., 2008b). However, to account for the occasional use of these areas, and to eliminate any artificial edge effects in the landscape resistance analyses, we expanded each population area by one kilometre outside of the limits defined above, except where coastline formed the edge. We quantified the landscape characteristics of each genetic population within the areas defined above using Geospatial Modelling Environment (www.spatialecology.com).

For each of the defined genetic populations, we constructed landscape resistance surfaces representing hypothesised effects of landscape variables on bobcat gene flow for fine-scale landscape genetic analyses. For each resistance surface, spatial data sources, resistance surface parameterisation, and ecological justifications are presented in Table 3.1. Landscape variables hypothesised to be positively related to gene flow were topographic roughness, vegetation density, and streams, while urbanisation and roads were hypothesised to be negatively related to gene flow. Each categorical resistance surface (roads, rivers) was weighted so that the minimum resistance = 10, and the maximum resistance = 100. Resistance surfaces representing continuous variables (topographic roughness, vegetation, urbanisation) were weighted on a scale of 1 to 100. We also generated an undifferentiated resistance surface for each population, representing isolation by distance (IBD; null model). We resampled all population-specific resistance surfaces to ensure a consistent resolution of 30 m by 30 m. For the region-wide landscape genetic analysis, we constructed resistance surfaces encompassing the entire study area, which we resampled to a resolution of 60 m by 60 m due to computational constraints. Preliminary tests across a smaller area produced highly similar results among 30 m and 60 m resolutions. Circuitscape 4.0.5 (McRae et al., 2008) was used to model connectivity between individuals along each resistance surface to

produce pairwise matrices of hypothesised landscape resistance to gene flow among individuals. Circuitscape was run for each population individually, as well as across the entire region at once.

Associations between landscape resistance matrices and pairwise genetic distances were tested using linear mixed effect models incorporating a maximum likelihood population effects (MLPE) approach (Clarke et al., 2002; Van Strien et al., 2012) using the *lme4* package in R. This method incorporates a random effect structure that accounts for the non-independence among pairwise data, and has been shown recently to outperform other model selection methods for landscape genetics (Shirk et al., 2018). Prior to fitting models, matrices of D_{ps} were log-transformed to satisfy normality assumptions and all dependent and independent variables were rescaled to units of standard deviation and a mean of zero. Ten models were fitted per population, including four multivariable models and six single-variable models (Table 3.2). Multivariable models were built according to general hypotheses about how gene flow might be influenced by landscape. For example, the 'VEGETATION' model explored the hypothesis that the presence of vegetation is driving gene flow patterns and included as fixed effects vegetation density, streams (accounting for increased riparian vegetation density), and impervious surfaces (urban areas contain less vegetation). Other multivariable models were 'FULL' (all fixed effects included), 'HUMAN' (anthropogenic factors are driving gene flow), and 'LINEAR' (linear features, i.e., roads and streams, act as barriers to or corridors for gene flow). Multicollinearity among fixed effects were assessed for each multivariable model by calculating the variance inflation factor (VIF). Variables with VIF>10 were considered highly correlated and were excluded from final models.

Table 3.1. Landscape resistance variables tested.

Landscape variable	Description	Data source and original raster resolution	Landscape resistance hypothesis	Ecological justification
Isolation by distance (IBD)	Null model representing an isolation by distance effect	N/A	Homogeneous resistance surface; every cell has a resistance of 1	Genetic differentiation increases with geographical distance (Wright, 1943)
Impervious surface (IMPERV)	Measure of urban density, taken as the percentage landcover of impervious surfaces (e.g. buildings, concrete) per cell	2011 National Land Cover Database (mrlc.gov/nlcd2011.php; Homer et al., 2015) 30 m	Resistance increases with increasing % impervious surface. Linear relationship 1-100	Bobcats have been shown to avoid urban areas (Riley et al., 2003; Lyren et al., 2008b; Ordeñana et al., 2010)
Roads (ROAD)	All major and minor roads	OpenStreetMap (http://download.geofabrik.de/north-america/us/california.html) 30 m	All cells within 30m of a road have high resistance (100), all other cells have low resistance (10)	Bobcats prefer habitat containing fewer roads (Poessel et al., 2014)
Streams (STREAM)	All linear surface water features, including perennial, intermittent, and ephemeral streams, canals, and artificial channels	National Hydrography Dataset (nhd.usgs.gov) 30 m	All cells with 50m of a stream have low resistance (10), cells within 75m of a stream have moderate resistance (50), cells within 100m have moderate-high resistance (75), and all other cells have high resistance (100)	Bobcats favour riparian corridors for dispersal (Hilty & Merenlender, 2004; Jennings and Zeller, 2017)
Topographic roughness (ROUGH)	Unitless index representing the amount of elevation variation with a 3-by-3 cell moving window, calculated using Geomorphometry & Gradient Metrics Toolbox (Evans et al., 2014)	National Elevation Dataset (lta.cr.usgs.gov/ned) 10 m	Cells with higher topographic roughness have lower resistance; transformed from raw values to resistances from 1-100 using the 'MSSmall' function†	Carnivores are known to use topographic features as movement corridors (Lee & Vaughan, 2003; Dickson & Beier, 2007)
Vegetation density (VEG)	Enhanced Vegetation Index calculated from chlorophyll reflectance satellite imagery, measured in 2016	Moderate Resolution Index Imaging Spectroradiometer (modis.gsfc.nasa.gov) 250 m	Cells with higher vegetation density have lower resistance; transformed from raw values to resistances from 1-100 using the 'Small' function†	Bobcats more often occur in natural, vegetated habitats (Ordeñana et al., 2010)

† *Transformation functions implemented using the "Rescale by Function" tool in ArcGIS (ESRI)*

Table 3.2. Models for comparing individual pairwise genetic distances (Dps) to landscape resistance distances in each population.

Model name	Model parameters
FULL	D_{ps} ~ ROUGH + VEG + IMPERV + ROAD + STREAM
VEGETATION	D_{ps} ~ VEG + IMPERV + STREAM
HUMAN	D_{ps} ~ IMPERV + ROAD + ROUGH
LINEAR	D_{ps} ~ ROAD + STREAM
UNIV.IMP	D_{ps} ~ IMPERV
UNIV.VEG	D_{ps} ~ VEG
UNIV.RD	D_{ps} ~ ROAD
UNIV.STRM	D_{ps} ~ STREAM
UNIV.TR	D_{ps} ~ ROUGH
IBD	D_{ps} ~ IBD

MLPE models were initially fitted and evaluated for each population using the Bayesian Information Criterion (BIC), which Row et al. (2017) found to outperform R^2 for ranking models. All models with a ΔBIC<5 were considered candidates. Marginal R^2 were reported for descriptive purposes (calculated using the MuMIn package; Bartoń, 2014). Well-supported candidate models were then refitted using restricted maximum likelihood (REML) for unbiased estimation of beta coefficients (Clarke et al., 2002; Van Strien et al., 2012). Row et al. (2017) found that inclusion of an undifferentiated resistance variable representing isolation by distance (IBD) was effective in factoring out the effect of distance in MLPE models and reduced the likelihood of Type I error in estimating landscape resistance variable significance. Therefore, we also included an IBD fixed effect to all candidate models to improve accuracy of beta coefficient estimates. We calculated upper and lower 95% confidence intervals (CIs) of beta coefficients for all variables in REML-refitted candidate models. Model averaging of variable beta coefficients was conducted for each population using BIC evidence weights. Variables that had positive beta coefficients with CIs that did not overlap zero were considered to have a significant effect. Using this approach, landscape

resistance variables with significant negative beta coefficients typically indicate a non-true relationship (Row et al., 2017) and were thus interpreted here as non-significant.

3.4 Results

3.4.1 Genotyping and data filtering

Minimal variation in SNP numbers and error rates (locus, allele, and SNP) was observed among replicate samples due to changes in the *denovo_map.pl* parameter settings. Nonetheless, optimal parameter settings were chosen according to these measures, resulting in an average locus error rates of 0.289, an average allele error rate of 0.066, and an average SNP error rate of 0.027 prior to filtering of the SNP matrix. Initial processing of raw sequencing data in Stacks using optimal parameter settings resulted in a matrix of 141,705 SNPs among 293 individuals. Following filtering of individuals and loci for missing data, and of loci for low minor allele frequencies, outlier loci, and exclusion of SNPs at read positions 94 and 95, a final matrix of 13,520 SNPs among 278 individuals remained for analysis of population genomic structure (see Appendix Table A3.1 for detailed filtering results).

3.4.2 Population genomic structure

Analyses of population genomic structure using *fastStructure* indicated distinct genetic clusters northwest and southeast of Los Angeles, which were further structured with respect to certain major roads. For the northwest of Los Angeles group (Figure 3.1a), K = 2–3 was identified as optimal, with cross-validation error being lowest at K = 3 (Figure 3.2a). We chose K = 3 as the most plausible number of genetic clusters based on these results as well as consistency with existing knowledge of major geographical barriers to bobcat movement in

this area. These results indicated one genetic population located north of California State Highway 101 (this population subsequently referred to here as North-101), and two other populations that were both located south of Highway 101: one west of Interstate 405 (South-101) and one east of Interstate 405 (East-405).

For the southeast of Los Angeles group (Figure 3.1b, c), K = 1–3 was identified as optimal, with cross-validation error being lowest at K = 4 (Figure 3.2b). These analyses had difficulty consistently resolving genetic groups in this region, with the only repeated delineation being among individuals sampled east and west of Interstate 5 (I-5). Within these two groups, some substructure was identified, but not consistently among runs at different K values. Thus, we determined that the southeast of Los Angeles group contained two genetic populations. These two clusters consisted of individuals sampled from the San Joaquin Hills area west of I-5 (this population and the area it occupies are subsequently referred to as West-5), and one population east of I-5 (East-5) consisting of individuals across a large area in the Santa Ana Mountains south of Los Angeles and north of San Diego (Figure 3.1b). For quantifying rates of migration and admixture among West-5 and East-5, our results for K = 3 appeared to be most informative, so we identified among-population migrants using this model. However, close scrutiny of K = 2 and K = 4 reveals support for the same migrant individuals. Substructure within the West-5 population detected with K = 3 did not appear to correspond to any spatial or temporal pattern. Substructure within East-5 was only detected at K = 4, with clusters spatially organised in a roughly north-to-south pattern. However, these clusters overlapped spatially and did not appear to correspond to any known geographical barriers, potentially indicating a clinal pattern of genetic variation (Appendix 1 Figure A3.1, A3.2) driven by isolation-by-resistance relationships. Please see Appendix 2 for an additional analysis that corroborates our *fastStructure* findings.

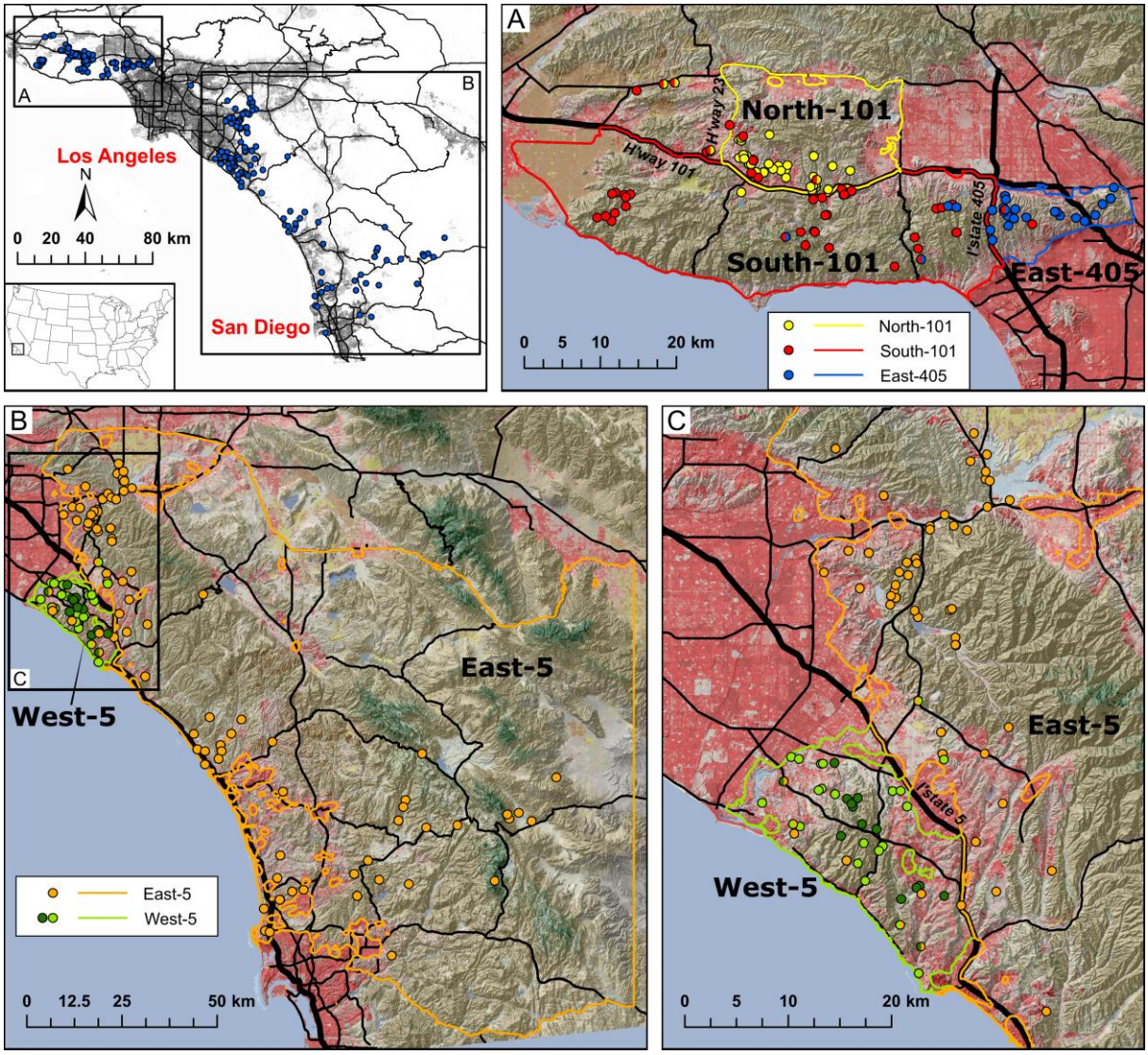

Figure 3.1. fastStructure analysis indicates five bobcat populations genetically isolated by geographic barriers in coastal Southern California, USA. All sampled individuals are shown as blue circles in the region-wide black-and-white map, with shading indicating urban development. Insets A–C indicate populations defined for landscape genomic analyses. Population spatial boundaries are indicated as solid coloured lines, and sample locations are coloured according to individual population assignment based on fastStructure analysis at K = 3 for both northwest and southeast of Los Angeles (Figure 3.2). Inset A shows populations northwest of Los Angeles, with East-405 (n = 25) indicated in blue, South-101 (n = 40) indicated in red, and North-101 (n = 50) indicated in yellow. Insets B and C show populations southeast of Los Angeles, with East-5 (n = 100) indicated in orange and West-5 (n = 42) indicated in light green and dark green. Individuals with greater than 25% admixture are shown with multiple colours. Highways are shown in all maps as black lines, with primary barriers to host gene flow (highway 101, Interstate 405, and Interstate 5) indicated by thick black lines. Topography and land cover are also shown in colour maps, indicating urbanisation in pink/red, forests in green, grasslands and scrub in tan, and agricultural areas in brown.

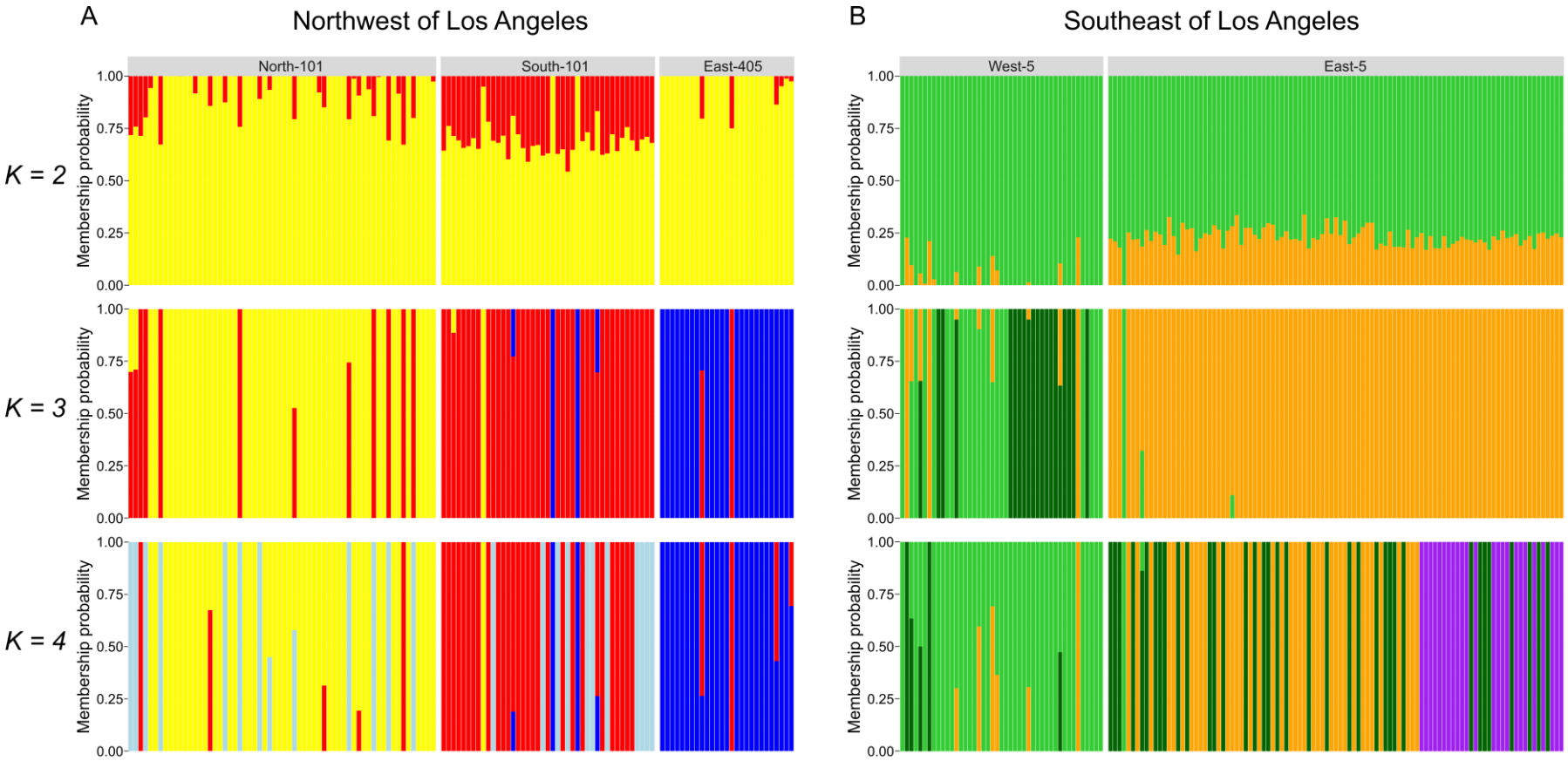

Figure 3.2. fastStructure analysis indicates population genetic structure among bobcats sampled from (A) northwest of Los Angeles and (B) southeast of Los Angeles, for K = 2, 3, and 4. Individuals are organised along the x-axis according to distance from the boundary of the nearest neighbouring population, with individuals from North-101 in order of decreasing distance from South-101, individuals from South-101 in order of increasing distance from North-101, and individuals from East-405 in order of increasing distance from South-101. Individuals from West-5 are organised in order of decreasing distance from East-5, and individuals from East-5 are organised in order of increasing distance from West-5.

We detected 21 migrant individuals across the study area. *fastStructure* assigned 12 individuals located north of Highway 101 to the South-101 population, including a distinct spatial group of six individuals located west of Highway 23. Three of these South-101-assigned individuals west of Highway 23 showed substantial admixture with the North-101 population, indicating that the area north of Highway 101 and west of Highway 23 is likely receiving immigrants from both North-101 and South-101. Due to the uncertainty in assigning individuals from the area west of Highway 23, this area was excluded from subsequent population-specific analyses. The South-101 population area contained one individual that assigned to the North-101 population, and two individuals that assigned to East-405. The East-405 population area contained two individuals that assigned to the South-101 population. We identified less migration among the two populations located southeast of Los Angeles, with one individual located east of Interstate 5 assigned to West-5, and three individuals located west of I-5 assigned to East-5. Pairwise F_{ST} values were significant among all populations and ranged between 0.041–0.150 (Table 3.3), supporting the genetic clusters identified using *fastStructure.*

Table 3.3. Pairwise FST among each population pair (95% confidence intervals shown in parentheses), indicating varying degrees of genetic differentiation among populations.

	East-5	West-5	East-405	South-101
West-5	0.046 (0.039–0.055)	-	-	-
East-405	0.100 (0.089–0.120)	0.150 (0.130–0.170)	-	-
South-101	0.041 (0.036–0.046)	0.077 (0.070–0.086)	0.077 (0.065–0.093)	-
North-101	0.061 (0.055–0.067)	0.100 (0.092–0.110)	0.110 (0.100–0.140)	0.051 (0.044–0.060)

3.4.3 Genomic diversity and demographic parameters

Following removal of migrants, 257 individuals remained for population and landscape genomic analyses. The population East-405 was the least genetically diverse, having the lowest measures of allelic richness and nucleotide diversity, with these measures also relatively low in West-5 (Table 3.4). The largest populations by geographical area (East-5 and South-101) had the highest measures of allelic richness and nucleotide diversity. Effective population sizes were generally congruent with the genetic diversity measures, with populations with higher genetic diversity having higher effective population sizes.

Table 3.4. Genetic diversity statistics for each population with the number of genotyped, non-migrant individuals used for calculating these statistics (n), allelic richness (Ar), observed heterozygosity (H_{obs}), expected heterozygosity (H_{exp}), inbreeding coefficient (F_{is}), and effective population size (Ne). The greatest genetic diversity and effective population sizes were observed in East-5 and South-101, while East-405 contained the lowest diversity and smallest effective population size.

Population	n	Ar	H_{obs}	H_{exp}	F_{is}	N_e
North-101	50	1.72	0.1546	0.1861	0.1181	22.9 (16.4-33.6)
South-101	40	1.78	0.1661	0.1942	0.1105	90.3 (46.3-504.2)
East-405	25	1.59	0.1425	0.1652	0.0797	12.8 (6-35.1)
West-5	42	1.68	0.1664	0.1835	0.0701	18.9 (14.4-25.5)
East-5	100	1.82	0.1531	0.1966	0.1951	150.3 (100.5-271.8)

3.4.4 Landscape genomics

Habitat areas for each population varied in size as well as their landscape composition (Table 3.5). The population with the largest area was East-5 (15,067 km^2), with all other populations below 1,000 km^2. East-5 was also the least urbanised and had the fewest roads, with the highest degrees of urbanisation and road density being in populations with the smallest habitat areas, East-405 and West-5. Stream density was relatively consistent among

populations, except for East-405 which had comparatively few streams. Topographic roughness was highest in the South-101 population, which encompassed the Santa Monica Mountains, and was lowest in West-5.

No collinearity was detected among predictors, except among the resistance distances for impervious surfaces and roads in the East-405 population; we thus excluded impervious surfaces from all multivariable models for East-405 (although we retained it for the single-variable model). For the region-wide landscape resistance analysis, linear mixed effects models with MLPE returned only the full model as a candidate according to BIC (Table 3.6). Within this model, we found significant effects of topographic roughness and impervious surfaces with roughness positively associated and impervious surfaces negatively associated with gene flow, but no significant effect of IBD, vegetation, roads, or streams (Table 3.6, Figure 3.3). The population inhabiting the largest area, East-5, showed significant effects of vegetation, impervious surfaces, and streams, but no effect of IBD, indicating that vegetation and streams were positively associated with gene flow, and impervious surfaces were negatively associated with gene flow.

Table 3.5. Landscape characteristics for each genetic population ± standard deviation, indicating substantial variation among populations. Linear features (roads and streams) are given as average length of features per square kilometre, with all other features given as the average raster cell value across the population area. Urbanisation (impervious surface) are actual percentage values (i.e., 100% urbanisation means that the impervious surfaces cover 100% of a raster cell), while vegetation density and topographic roughness are percentages relative to the highest and lowest values of each variable across the whole region.

Population	Total area (km^2)	Vegetation density (%)	Urbanisation (%)	Roads (km/km^2)	Streams (km/km^2)	Topographic roughness (%)	Elevation (m)
North-101	406	38.2 ± 7.27	16.9 ± 23.0	5.55	1.57	2.35 ± 2.49	348 ± 100
South-101	959	49.3 ± 12.2	9.45 ± 19.1	3.60	1.53	3.74 ± 4.01	280 ± 174
East-405	185	47.1 ± 11.5	28.8 ± 29.1	9.41	0.53	2.93 ± 3.37	225 ± 89.6
West-5	340	43.4 ± 10.2	27.7 ± 27.8	9.22	1.97	1.83 ± 2.11	98.1 ± 67.7
East-5†	15,067	38.4 ± 15.9	6.68 ± 17.5	2.43	1.52	2.62 ± 3.20	632 ± 445

† Landscape characteristics for this population are estimates only as it has a high degree of openness to external source populations without a clearly defined geographical boundary (e.g., major highways known to be acting as barriers).

Table 3.6. Linear mixed effect modelling with maximum-likelihood population effects indicates significant effects of landscape resistance variables on gene flow within genetic populations as well as across the entire region. Candidate models are listed according to ΔBIC (up to a maximum of 5) calculated from initial fitting of models without restricted maximum likelihood (REML) and without the isolation by distance (IBD) fixed effect. Variance inflation factors (VIF) and marginal R^2 values reported are from these initial models. 95% confidence intervals for parameter beta coefficients (β) were calculated from refitting of candidate models with REML and IBD fixed effect included for enhanced accuracy and reduction of Type I error. 95% confidence intervals greater than and not overlapping zero indicate parameter significance and are shown in bold with an asterisk. See table 2 for model and parameter descriptions.

Population	Candidate models	ΔBIC	mR^2	Parameter	β	95% CI		VIF
						Upper	Lower	
Region	FULL	0.00	0.64	Intercept	0.001	-0.385	0.387	
				IBD	-0.300	-0.557	-0.043	
				ROUGH*	**0.923**	**0.675**	**1.172**	6.07
				VEG	-0.219	-0.334	-0.103	2.88
				IMPERV*	**2.333**	**2.172**	**2.493**	2.90
				ROAD	-0.223	-0.404	-0.042	4.96
				STREAM	-0.435	-0.568	-0.302	3.48
North-101	UNIV.ROAD	0.00	0.058	Intercept	0.004	-0.236	0.244	
				IBD	0.021	-0.188	0.231	
				ROAD	0.201	-0.009	0.411	
	UNIV.TR	2.62	0.066	Intercept	0.005	-0.253	0.262	
				IBD	0.059	-0.290	0.407	
				ROUGH	0.177	-0.210	0.565	
	IBD	3.31	0.053	Intercept	0.005	-0.250	0.260	
				IBD*	**0.216**	**0.161**	**0.272**	
	UNIV.RIV	4.37	0.124	Intercept	0.005	-0.243	0.253	
				IBD	0.124	-0.012	0.260	
				STREAM	0.162	-0.057	0.380	

	LINEAR	4.72	0.068	Intercept	0.004	-0.225	0.233	
				IBD	-0.097	-0.335	0.141	
				ROAD*	**0.218**	**0.013**	**0.423**	1.25
				STREAM	0.177	-0.026	0.380	1.25
South-101	UNIV.IMP	0.00	0.075	Intercept	0.001	-0.371	0.374	
				IBD	-0.275	-0.666	0.117	
				IMPERV*	**0.514**	**0.116**	**0.913**	
	UNIV.RD	1.30	0.113	Intercept	0.001	-0.358	0.360	
				IBD	0.006	-0.245	0.256	
				ROAD	0.292	-0.033	0.616	
	UNIV.RIV	1.34	0.084	Intercept	0.000	-0.346	0.346	
				IBD	-0.140	-0.552	0.273	
				STREAM	0.396	-0.047	0.839	
	UNIV.TR	3.18	0.078	Intercept	0.001	-0.358	0.359	
				IBD	-0.228	-0.992	0.535	
				ROUGH	0.481	-0.325	1.286	
	IBD	3.96	0.071	Intercept	0.001	-0.336	0.337	
				IBD*	**0.226**	**0.177**	**0.275**	
East-405	UNIV.IMP	0.00	0.269	Intercept	0.006	-0.368	0.380	
				IBD	-0.635	-1.062	-0.208	
				IMPERV*	**1.090**	**0.664**	**1.515**	
West-5	IBD	0.00	0.125	Intercept	0.001	-0.258	0.260	
				IBD*	**0.345**	**0.265**	**0.425**	
East-5	VEGETATION	0.00	0.318	Intercept	0.000	-0.249	0.248	
				IBD	-0.069	-0.219	0.082	
				VEG*	**0.348**	**0.199**	**0.498**	1.29
				IMPERV*	**0.209**	**0.038**	**0.379**	1.11
				STREAM*	**0.294**	**0.124**	**0.463**	1.35

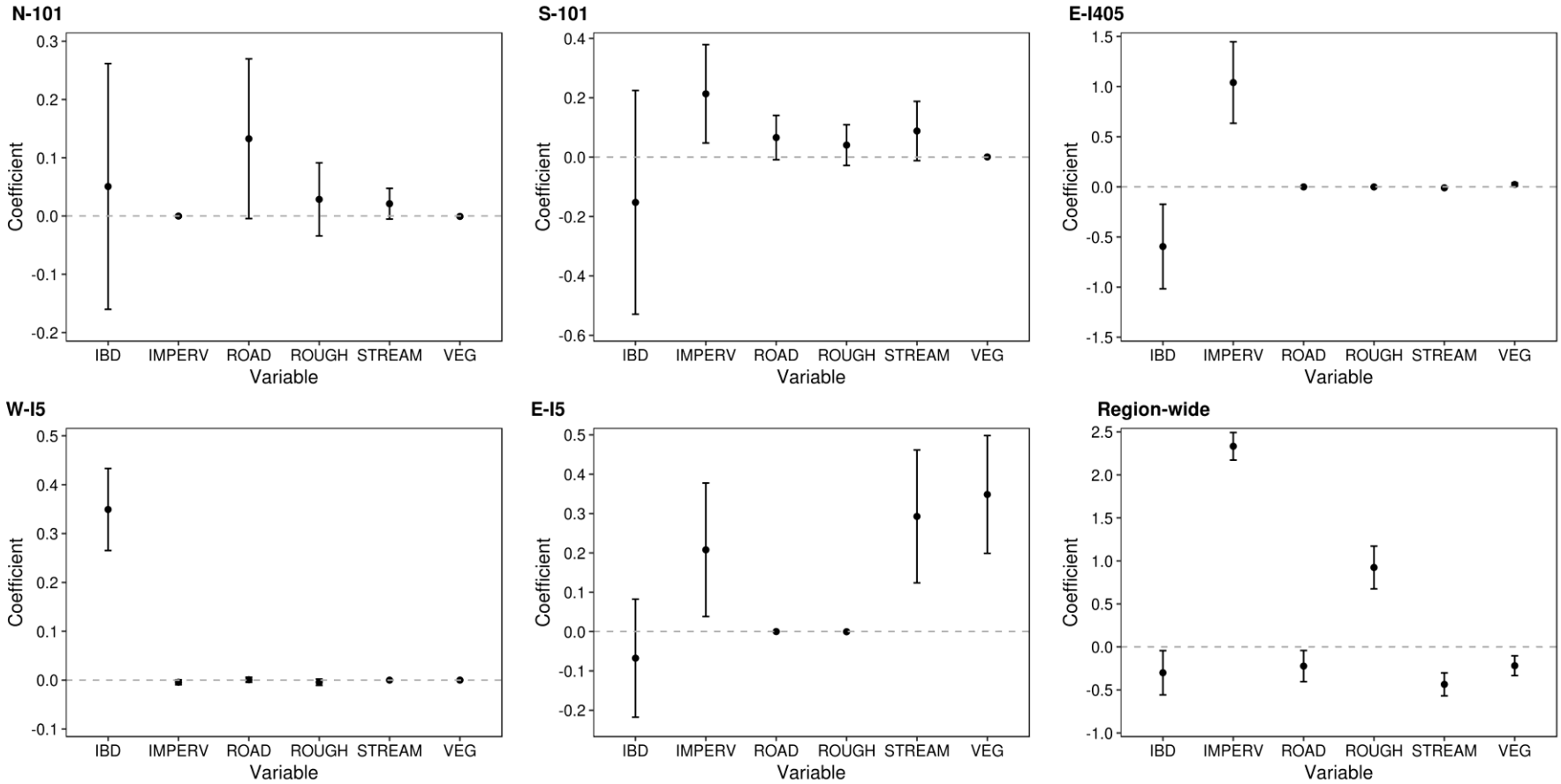

Figure 3.3. Average coefficients from MLPE candidate models indicate significant effects of landscape resistance variables on individual pairwise genetic distances among bobcats within five separate populations in coastal southern California, as well as among all individuals across the region. Mean coefficients are shown with upper and lower confidence intervals, weighted according to BIC model support. IBD = isolation by distance; IMPERV = impervious surfaces; ROAD = roads (major and minor) and highway links; ROUGH = topographic roughness; STREAM = all ephemeral, intermittent and perennial surface waterways; VEG = vegetation density.

Among the spatially smaller populations, generally fewer significant landscape effects on genetic distances were identified. We found strong effects of impervious surfaces on gene flow in both South-101 and East-405, which had relatively low densities of streams (Figure 3.4). Conversely, populations exhibiting evidence for streams being positively associated with gene flow (East-5, South-101, and North-101) had generally lower urbanisation and road density and had larger areas (Figure 3.4). However, two of these populations showed only near-significant support for streams (South-101 and North-101; Figure 3.3, Figure 3.4). There was near-significant support for a negative effect of roads on gene flow in North-101 and South-101 (Figure 3.3). IBD was the only supported predictor of pairwise genetic distances in West-5. Marginal R^2 values were highest in the region-wide candidate model (mR^2 = 0.640), East-5 (mR^2 = 0.318) and in East-405 (mR^2 = 0.269), and lowest in North-101 (mR^2 = 0.053–0.124), indicating substantial variation among populations in the ability of our models to explain individual genetic variation.

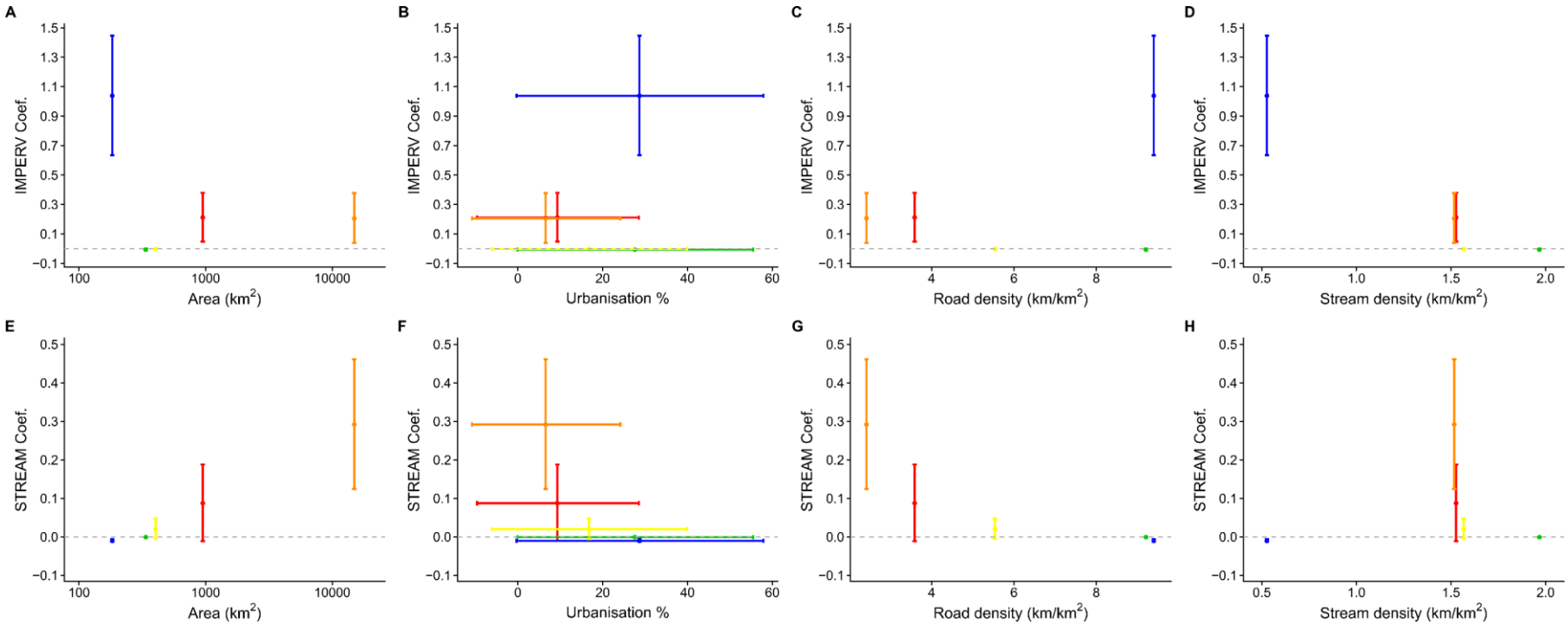

Figure 3.4. Trends in average beta coefficients for the effects of resistance distances representing impervious surfaces (A–D) and streams (E–H) on individual pairwise genetic distances in each population, relative to landscape characteristics of the area inhabited by each population. The effect of impervious surfaces on connectivity is generally greater in populations with fewer streams, while the effect of streams on connectivity is generally greater in populations encompassing a larger spatial area with lower urbanisation and lower road density. Vertical error bars represent 95% confidence intervals; horizontal error bars represent ± 1 standard deviation. Colours indicate study population, with East-405 shown in blue, South-101 shown in red, North-101 shown in yellow, East-5 shown in orange, and West-5 shown in green. Comparisons between all combinations of resistance distance coefficients and population area characteristics are shown in Appendix Figure A3.1.3.

3.5 Discussion

3.5.1 Effect of spatial scale on landscape genetic inference

Understanding how landscape features influence connectivity within habitat patches is important for identifying factors that maintain connectivity and to elucidate the impacts of habitat degradation. Our replicated landscape genomic analyses enabled a rigorous investigation into which landscape factors affect bobcat functional connectivity at both local (within populations) and regional (among populations) spatial scales. In most of our studied populations, a proportion of the variation in genetic distances among individuals was attributable to one or more of our landscape resistance hypotheses. However, the spatial scale over which these hypotheses were tested appeared to be an important factor in determining their level of support, with the greatest support appearing to be among analyses conducted over larger areas. This relationship with spatial scale was reflected in the proportions of genetic variation explained by our models, which explained up to 64% of genetic variation at the regional scale and 32% of variation in our spatially largest population (East-5), but generally explained less than 13% of variation within most of our populations encompassing smaller areas (with the exception of 27% in East-405, discussed below). The ratio of dispersal distance to study area size is an important factor determining the proportion of spatial genetic structure that is likely to be explained by landscape heterogeneity, with a greater proportion of variation explained when this ratio is lower. For example, in studies of species with very low dispersal distances compared to bobcats, landscape genetic models can explain upwards of 40% of genetic structure across study areas comparable in size to our smaller populations (Funk et al., 2005; Wang, 2009; Goldberg & Waits, 2010; Murphy et al., 2010b).

We found strong support for models containing multiple significant effects of landscape variables in both the region-wide analysis and East-5. In contrast, we observed few effects of landscape resistance in the four populations with the smallest habitat areas, with West-5 showing no support for any resistance hypothesis other than IBD, and only minimal statistical support for any resistance hypothesis in North-101. Nonetheless, we detected strong support for impervious surfaces restricting gene flow in South-101 and the population with the smallest habitat area, East-405. Sample size did not appear to be a major factor affecting inferential power among our populations. Despite seeing higher support across larger spatial areas for landscape factors influencing gene flow, the smaller areas to which some populations were confined did not preclude us from detecting landscape genetic signals where connectivity was strongly impacted.

For the spatial scales at which habitat fragmentation is occurring in coastal southern California, bobcats are excellent indicators of functional connectivity. This is largely due to their intermediate sensitivity to anthropogenic disturbance and reliance on large, connected areas of natural habitat (Crooks, 2002). Hard barriers such as highways or tracts of urban development can have a substantial and highly detectable effect on bobcat gene flow (Riley et al., 2006; Lee et al., 2012; Serieys et al., 2015). However, within fully natural or altered natural areas, our results demonstrate that the high vagility and generalised habitat use of bobcats (Ordeñana et al., 2010; Riley et al., 2010) means that the effects of specific landscape factors on functional connectivity are often subtle and their detectability influenced by spatial scale, even with the high precision afforded by large genomic datasets.

3.5.2 Patterns of population structure

Northwest of Los Angeles, California State Highway 101 and Interstate 405 formed significant barriers to gene flow and produced three genetically distinct populations, with occasional migration across highway barriers, particularly from south-to-north over Highway 101. This is generally consistent with previous research (Riley et al., 2006; Serieys et al., 2015). Similarly, southeast of Los Angeles, the genetic break either side of Interstate 5 is consistent with that observed from microsatellite data, again with low rates of migration (Lee et al., 2012; Ruell et al., 2012). Large SNP datasets hold great utility over microsatellites where precise measurements of subtle genetic variation are required, such as for individual-based landscape genetics in highly mobile species such as bobcats. However, our results suggest that microsatellites may be sufficient for studies assessing patterns of population genetic structure where delineations among populations are relatively clear (such as highways forming a major impediment to connectivity). Ultimately, in coastal southern California, major freeways represent the most acute barriers to connectivity among bobcat populations, necessitating further development of corridors to better facilitate crossing.

A notable difference in our results compared to microsatellite studies was that we did not identify temporally partitioned genetic structure north of 101 with respect to a disease outbreak, reported by Serieys et al. (2015). This notoedric mange epizootic occurred in the North-101 population between 2002 and 2005, resulting in dramatic reductions in annual survival rates over this period (Riley et al., 2007). Serieys et al. (2015) found two temporal genetic clusters within North-101 that corresponded to before and after 2005, and found that differentiation between pre- and post-epizootic populations was greater than that between populations separated by major highway barriers. However, although the present study used a subset of the same individuals (pre- and post-epizootic), no evidence of this genetic structure

was found in our *fastStructure* analysis of SNP data. Additional *fastStructure* runs focusing only on samples from North-101 did not identify any spatial or temporal clustering for $K > 1$. It is unclear precisely why we did not detect this pattern. However, microsatellites generally have a greater number of alleles per locus than generally bi-allelic SNPs, including a higher number of rare alleles, which may have been lost from the population following the bottleneck associated with the mange outbreak. The comparatively greater loss of allelic diversity among the microsatellite loci compared to the SNPs may have resulted in a stronger signal of pre- vs. post-epizootic genetic differentiation.

Serieys et al. (2015) proposed that the pre- and post-epizootic differentiation they detected was more likely to have occurred due to genetic drift within a small population of survivors originating from North-101, rather than a recolonisation of this area by migrants from other populations. Our results support this conclusion; specifically, a post-epizootic immigration explanation was not supported by our data because, despite the large number of migrants we detected from South-101, many of these migrant individuals were captured prior to the mange outbreak. *fastStructure* results for $K = 4$ from this area indicated a possible additional cluster (indicated in light blue in Figure 3.2A) composed roughly equally of individuals captured in North-101 and South-101, also suggesting that the strongest substructure within North-101 is driven by migrants from South-101, and not another unsampled source population, such as from the Santa Susana mountains to the north. If migrants of different genetic origin were detected, we would likely see a novel cluster in North-101 that is minimally shared with the other populations sampled.

3.5.3 Implications for functional connectivity

Urban development containing impervious surfaces was the most frequently identified factor impacting connectivity, having a negative effect on gene flow. While bobcats generally occur more often in natural areas (Ordeñana et al., 2010), some telemetry studies show that bobcats will cross and sometimes utilise urban development separating habitat fragments, particularly at night (Tigas et al., 2002; Riley et al., 2003). Our results indicate that despite these movements, urban development does constrain gene flow even within areas that consist largely of natural habitat (e.g., East-5, South-101). The strongest effect of impervious surfaces appeared to be within our region-wide analysis due to urban development and major highways constraining gene flow among populations. At this broad scale, bobcat gene flow also favoured areas that were topographically rough, likely in part because such terrain precludes intensive urban development. This pattern is obvious in coastal southern California, where the largest tracts of urban development are situated in the flat regions of the Los Angeles Basin, coastal Orange County, and the coastal plain of San Diego County, and bobcat habitat generally is restricted to the surrounding Santa Monica Mountains, San Joaquin Hills, and eastern Peninsular Ranges.

Although we observed the effects of impervious surfaces in populations surrounded by relatively low urban development, we did not observe this pattern in some populations with moderate and high amounts of urban development (North-101 and West-5). As we outlined above, a lack of effect in these populations may be due to low detectability of landscape genetic signals at small spatial scales; however, variation among populations in dispersal behaviour may be a contributing factor as well. Studies in other regions with different patterns of urban development have found that bobcats more strongly avoid urban areas in landscapes that are less fragmented by urbanisation overall (Riley, 2006). In coastal

southern California, bobcats in areas with relatively broad tracts of natural habitat may be better able to avoid urban areas through greater availability of alternative routes. Additionally, a home-range pileup effect has been observed previously in this region whereby territories adjacent to urban and highway barriers tend to be smaller and more densely distributed (Riley et al., 2006). As a result, juveniles from more urbanised areas may be required to disperse greater distances through potentially less suitable habitat (e.g., areas containing more impervious surface) to find territory. Nonetheless, a strong negative effect of impervious surfaces on gene flow was detected in our smallest and most heavily urbanised population (East-405). For such a small patch of habitat, it is concerning that the effect of impervious surfaces explained such a high proportion of the among-individual genetic variation within this population (mR^2 = 0.269). Further urban development in this area has a risk of producing smaller, more isolated habitat patches that may be insufficient to support viable bobcat populations, resulting in localised extinction. The area east of I-405 no longer supports a viable population of mountain lions due to urban habitat fragmentation (Riley et al., 2014), and our results suggest that bobcats are at risk of a similar fate.

Streams and vegetation had significant effects on gene flow in only our largest and least urbanised population, East-5. In San Diego County, which encompasses a substantial part of our East-5 population, previous work has found associations between bobcat occurrence and water availability (Markovchick-Nicholls et al., 2008) as well as evidence for use of streams as dispersal corridors (Jennings & Zeller, 2017). In addition, telemetry has indicated bobcats in the San Joaquin Hills utilise riparian corridors, particularly as a means of traversing roads and urban areas (Lyren et al., 2008b), while camera surveys indicate bobcat use of riparian corridors in agricultural areas in northern California (Hilty & Merenlender, 2004). In contrast, our results suggest no overall effect of streams on gene flow in our San

Joaquin Hills population (West-5), nor in our most heavily urbanised population, East-405. Again, such an effect may not be detectable at the relatively small spatial scales of these populations, and indeed, there was a weakly supported positive effect of streams in North-101 and South-101. Populations experiencing less urbanisation and fewer roads appeared more likely to exhibit a positive effect of streams on gene flow (East-5 and South-101), while the strongest negative effect of urbanisation on gene flow was in the population containing the fewest streams (East-405). One potential consequence of urbanisation is the diversion of surface streams into channels and pipelines. This is particularly evident in the East-405 habitat area, where the lower areas of many catchments contain suburban development with pipelines running underneath. The resulting loss of streams as usable habitat may explain why East-405 exhibited the highest effect of impervious surface in constraining gene flow (along with no positive effect of streams).

3.5.4 Concluding remarks

Our study demonstrates the value of multiple study areas and spatial scales when investigating the effects of landscape features on functional connectivity. Because such effects may occur only at certain spatial scales or in particular landscape contexts, failure to consider multiple spatial scales can cause researchers to miss important relationships, or spuriously extrapolate such relationships to circumstances where they do not apply. Our results consistently indicate that urban development has a negative effect on landscape connectivity for bobcats, and we detected this effect at both local and region-wide spatial scales among multiple populations. This effect occurs even in areas that contain substantial amounts of natural habitat, as well as in areas where animals experience substantial

anthropogenic disturbance. We also found some evidence that streams have a positive effect on connectivity, particularly at larger spatial scales, and may help to mitigate the negative effects of urbanisation on connectivity. Continued alteration of natural areas through urban development and removal of riparian habitat is likely to place further pressure on carnivore populations in coastal southern California and may result in extirpation from some existing habitat fragments.

3.6 Appendix 1

Table A3.1. Steps for filtering of SNP matrix showing number of individuals/loci before and after each step. Final numbers of loci and individuals used for analysis are indicated in bold.

Remove	Before	After
1. Loci genotyped in <65% of individuals	141,705	44,819
2. Loci with minor allele frequency <1%	44,819	20,316
3. Loci at read position 95	20,316	19,955
4. Loci at read position 94	19,955	19,689
5. Loci with high variance in per-individual missing genotypes among libraries	19,689	13,649
6. Outlier loci (PCAdapt)	13,649	**13,520**
7. Individuals with >50% missing loci	293	282
8. Individuals missing location data	282	**278**

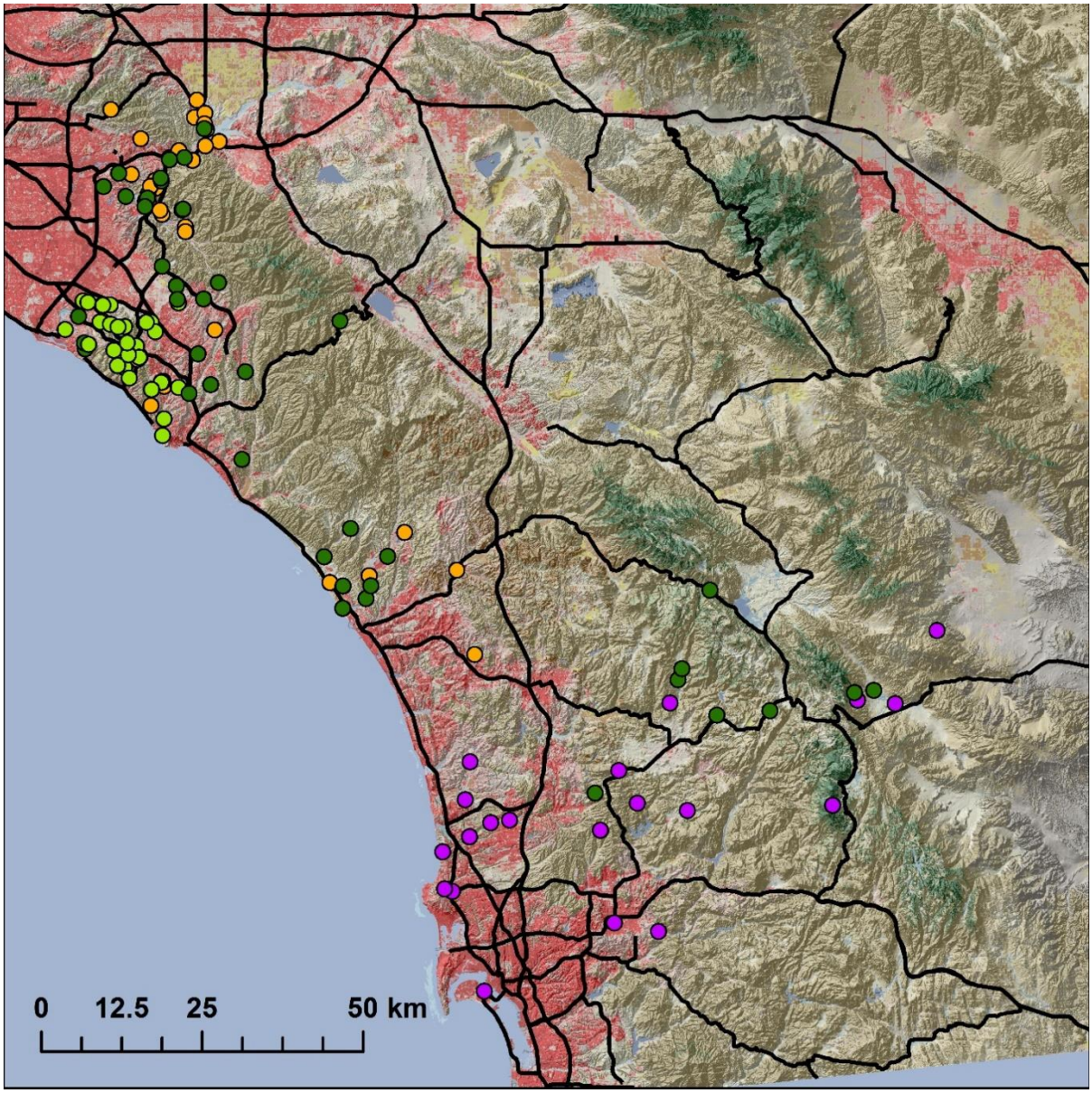

Figure A3.1.1. Clinal distribution of genetic clusters identified by fastStructure *analysis of bobcats southeast of Los Angeles using K = 4. Individuals from population West-5 largely form a single genetic cluster, while East-5 contains three overlapping clusters arranged in a north-to-south pattern that does not correspond to known geographic barriers.*

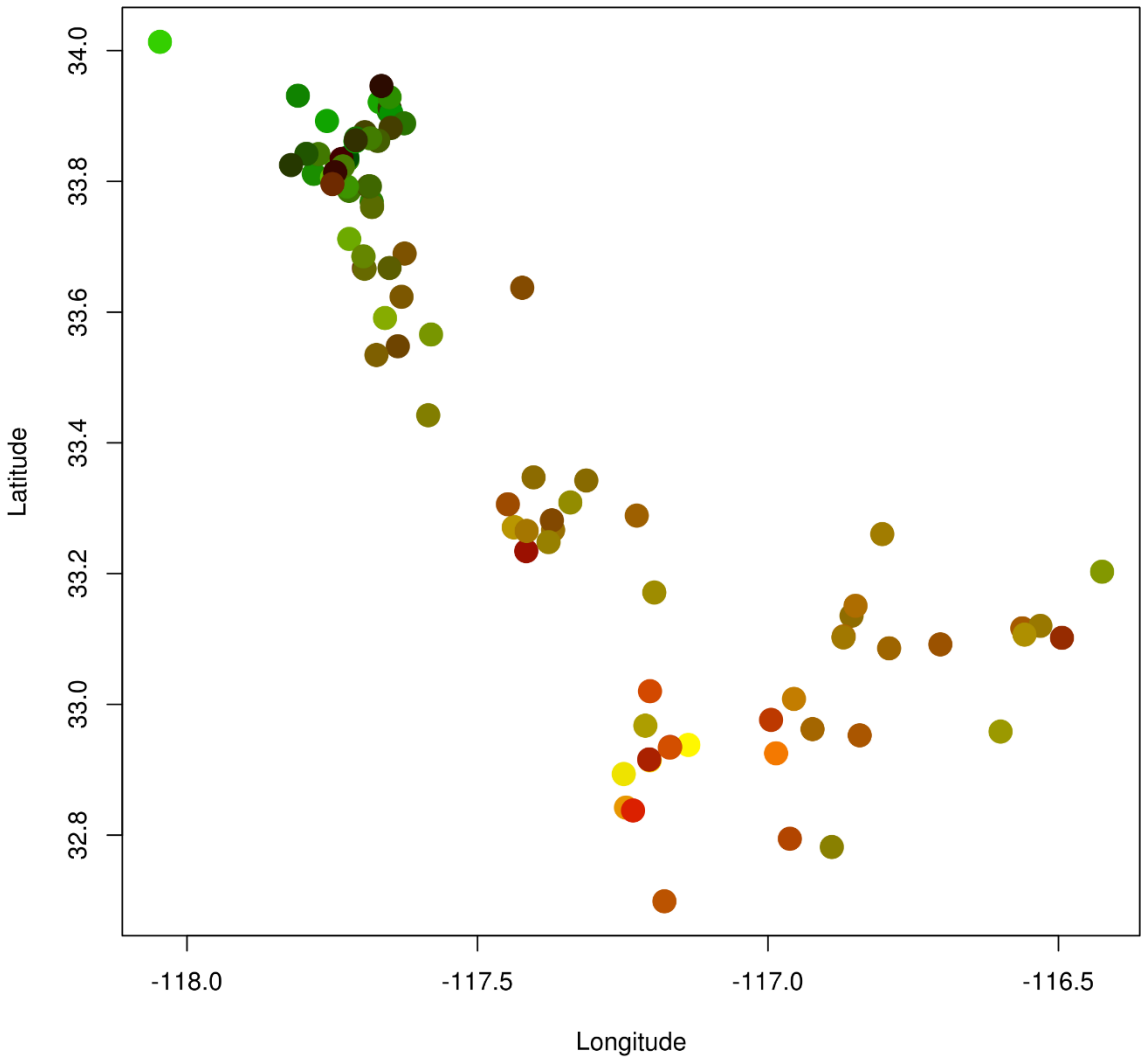

Figure A3.1.2. Results from spatial principal component analysis (sPCA; conducted using adegenet*) showing spatial genetic cline among bobcats in the East-5 population. Colours indicate first-axis sPCA scores representing spatial genetic variation among individuals. Markov Chain Monte-Carlo tests based on 10,000 permutations indicate significant (P << 0.05) global spatial genetic structure (positive spatial autocorrelation).*

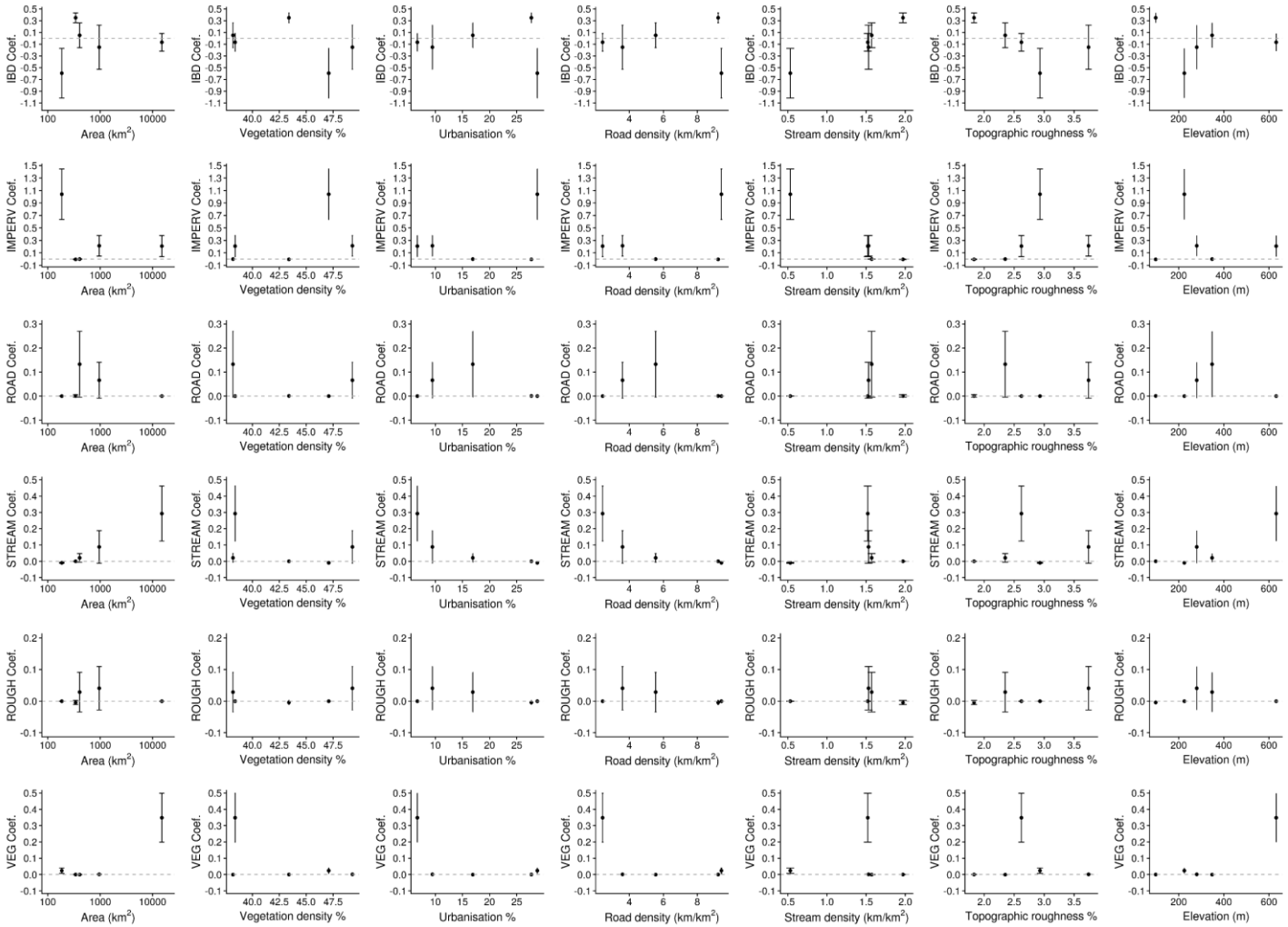

Figure A3.1.3. Average coefficients from MLPE candidate models for landscape resistance variables affecting individual pairwise genetic distances in each study area, relative to habitat characteristics of each study area. Vertical error bars represent 95% confidence intervals; horizontal error bars represent ±SD. IBD = isolation by distance; IMPERV = impervious surfaces; ROAD = roads (major and minor) and highway links; ROUGH = topographic roughness; STREAM = all ephemeral, intermittent and perennial surface waterways; VEG = vegetation density.

3.7 Appendix 2

3.7.1 Methods

To validate our *fastStructure* analyses, we performed an additonal analysis of population structure using an alternative, non-parametric approach: sparse nonnegative matrix factorization (SNMF; Frichot et al., 2014), implemented by the *LEA* package (Frichot & François, 2015) in R. Similarly to *fastStructure*, SNMF simultaneously estimates the number of genetic populations (K) and probabilistically assigns individuals to populations based on their multilocus genotypes, without any *a priori* population assignments. We conducted analyses for individuals sampled northwest and southeast of Los Angeles separately. SNMF was run for K = 1–10 with support for each estimated using a cross-entropy criterion calculated across 10 repetitions. Lower cross-entropy values indicate greater support for a given K.

3.7.2 Results and Discussion

Cross-entropy values indicate greatest support for K = 5 northwest of Los Angeles and K = 6 southeast of Los Angeles (Figure A3.2.1). However, upon examination of individual population assignments with respect to sample locations, it appears that these clusters do not necessarily reflect spatial population structure (Figure A3.2.2).

Northwest of Los Angeles, samples generally clustered according to the three spatial populations delineated by highway 101 and interstate 405 (Figure A3.2.2A). Evidence for three genetically and spatially distinct populations was observed at all K values, with

additional clusters for K > 3 tending to overlap spatially with one or more of the other clusters.

Similarly, southeast of Los Angeles, evidence for two genetically and spatially distinct populations separated by interstate 5 was observed at all K values (Figure A3.2.2B). Additional clusters for K > 2 tended to overlap spatially with one or more of the other clusters. However, at K > 5, some additional spatial structure was observed among individuals east of interstate 5, suggesting two highly admixed yet distinct groups separated north-to-south by no apparent geographical barrier, suggesting an isolation-by-distance pattern. Such a pattern was also supported by spatial principal components analysis (see 3.6 Appendix 1).

Overall, despite the high K values predicted by the cross-entropy criterion, our nonparametric SNMF analysis supports our *fastStructure* findings of three spatially and genetically distinct populations northwest of Los Angeles and two spatially and genetically distinct populations southeast of Los Angeles.

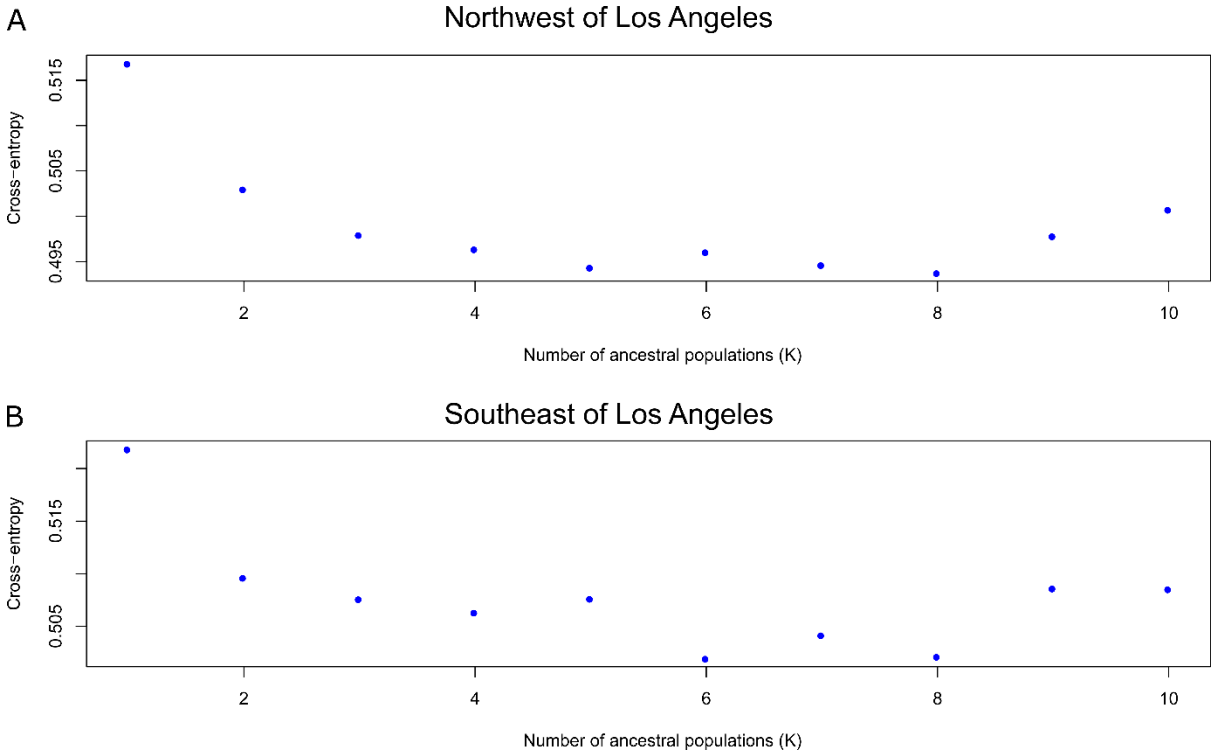

Figure A3.2.1. SNMF analysis indicates greatest support for K = 5 genetic populations northwest of Los Angeles (A) and K = 6 genetic populations southeast of Los Angeles based on a cross-entropy criterion calculated across 10 replications for each K.

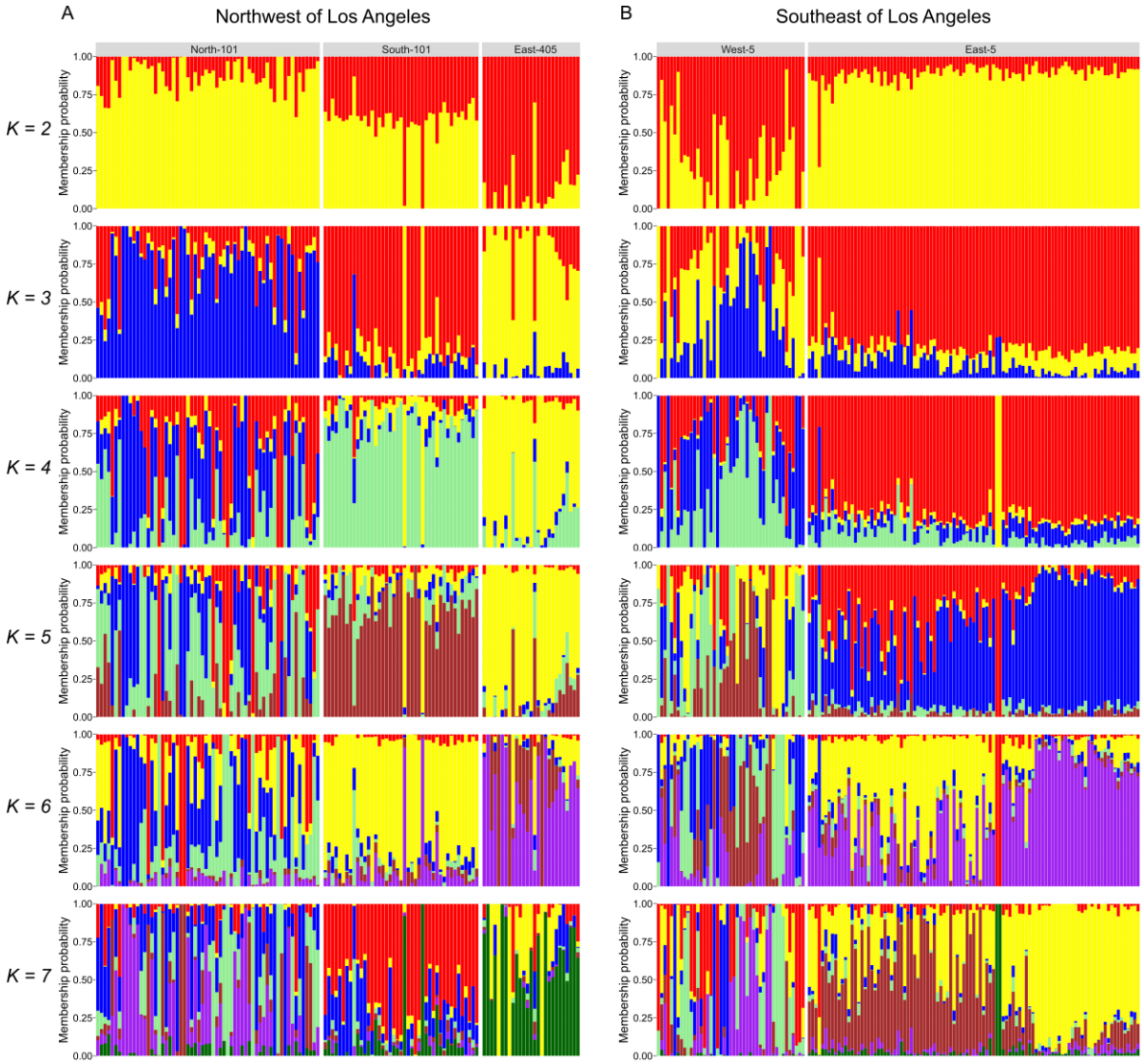

Figure A3.2.2. Per-individual population assignments performed by SNMF for K = 2–7 for individuals samples northwest (A) and southeast (B) of Los Angeles. Individuals are organised along the x-axis according to their location with respect to major highways. Individuals labelled North-101 were sampled north of and are arranged in order of decreasing distance from highway 101. Individuals labelled South-101 were sampled south of and are arranged in order of increasing distance from highway 101. Individuals labelled East-405 were sampled east of and are arranged in order of increasing distance from interstate 405. Individuals labelled West-5 were sampled west of and are arranged in order of decreasing distance from interstate 5, while individuals labelled East-5 were sampled east of and are arranged in order of decreasing distance from interstate 5.

Chapter 4

Does the virus cross the road? Viral phylogeographic patterns among urban bobcat populations

4.1 Abstract

Urban development has major impacts on connectivity among wildlife populations and is thus likely an important factor shaping pathogen transmission in wildlife. However, most investigations of wildlife diseases in urban areas focus on prevalence and infection risk rather than potential effects of urbanisation on transmission. Feline immunodeficiency virus (FIV) is a directly-transmitted retrovirus that infects many felid species and is a model for studying pathogen transmission at landscape scales. We reconstructed phylogenetic relationships among FIV_{Lru} isolates sampled from five bobcat populations in coastal southern California that appear isolated due to major highways and dense urban development. We found strong FIV_{Lru} phylogeographic structure among three host populations northwest of Los Angeles, largely coincident with host genetic structure. In contrast, relatively little FIV_{Lru} phylogeographic structure existed among two genetically-distinct host populations southeast of Los Angeles. Rates of FIV_{Lru} transfer among populations did not vary significantly, suggesting that the lack of phylogenetic structure southeast of Los Angeles may be a product of incomplete lineage sorting rather than frequent contemporary transmission among populations. Divergence dates among FIV_{Lru} lineages in several cases reflected historical urban growth and construction of major highways. Our results indicate that major barriers to host gene flow can also act as barriers to pathogen spread, suggesting potentially reduced susceptibility of these populations to outbreaks of novel pathogens.

4.2 Introduction

Urban development is one of the primary factors affecting wildlife worldwide (Ramalho & Hobbs, 2012). Urban influences can affect the abundance of wildlife populations and the frequency of intra- and inter-specific interactions (Crooks et al., 2010; Shochat et al., 2010), as well as reducing connectivity among populations (Bierwagen, 2007). In doing so, urbanisation may also alter patterns of pathogen transmission and spread (Becker et al., 2015; Fountain-Jones et al., 2017a). This can have consequences for the rate and severity of disease outbreaks, and the probability of pathogen spillover, into wildlife, humans and domestic animals (Faust et al., 2018). Most urban-disease research has examined how pathogen prevalence and infection risk is modified by urbanisation (Brearley et al., 2013), but relatively few studies investigate the impacts of landscape heterogeneity on patterns of pathogen transmission and spread. Many studies instead use host gene flow as a proxy for transmission risk (Kozakiewicz et al., 2018a), but this may be inappropriate in some cases because host and parasite genetic structure are often divergent (Mazé-Guilmo et al., 2016). Understanding the extent to which patterns of pathogen transmission are related to host population structure and the effects of habitat fragmentation on pathogen connectivity are important to predicting and managing wildlife disease in urban environments.

Bobcats (*Lynx rufus*) are regarded as indicator species for connectivity in urbanising regions given their intermediate sensitivity to anthropogenic disturbance and requirement for high quality, connected areas of habitat (Crooks, 2002). Thus, it follows that bobcat pathogens may similarly be used as a model for understanding disease connectivity in urbanising environments. Bobcat populations are susceptible to several highly pathogenic diseases, including feline leukaemia virus, feline panleukopenia virus, and notoedric mange (Sleeman et al., 2001; Riley et al., 2004, 2007; Carver et al., 2016), with patterns of spread potentially influenced by landscape connectivity as well as variation in environmentally- or

host-driven selective pressure (Real & Biek, 2007). Predicting how future outbreaks of such pathogens might spread among bobcat populations may be important from a conservation perspective (Smith et al., 2009), as well as for understanding the likelihood of spillover into domestic animal populations.

Feline immunodeficiency viruses (FIVs; the feline analogue of HIV) infect many felid species, including bobcats (FIV_{Lru}), and has characteristics that make it a suitable model for studying pathogen transmission, and also for inferring host movement (Biek et al., 2006; VandeWoude & Apetrei, 2006; Antunes et al., 2008; Kerr et al., 2018). FIVs are directly transmitted, generally species-specific, and form chronic, lifelong infections resulting in relatively uncomplicated transmission networks that remain largely intact owing to little overt pathology (VandeWoude & Apetrei, 2006). FIV proviral DNA can also be isolated from host blood or tissue and has high mutation rates, enabling phylogenetic reconstruction of recent transmission histories (Biek et al., 2003) and temporal estimates of divergence among viral populations (Drummond et al., 2002). Differences in FIV phylogeographic structure may indicate variation in rates of pathogen transmission and spread, which can arise through altered connectivity, or local adaptation of the pathogen to environmental conditions or host immunology whereby locally evolved viral variants are better able to infect local hosts than hosts from other populations (Kaltz & Shykoff, 1998; Real & Biek, 2007; Lion & Gandon, 2015). In addition, because increased isolation of host populations would likely decrease introduction of novel pathogen variants and place constraints on pathogen population sizes, we might expect reduced pathogen diversity in more isolated populations. Thus, FIVs can inform how habitat fragmentation and its impacts on host connectivity from urbanisation influence the spread of directly transmitted pathogens.

Several genetically distinct bobcat populations are evident northwest and southeast of Los Angeles, with major freeways and large urban areas forming the primary barriers to gene

flow (Riley et al., 2006; Lee et al., 2012; Ruell et al., 2012; Serieys et al., 2015). Intensifying urbanisation is also generally reducing connectivity within populations; however, riparian corridors may play a role in maintaining connectivity in urban areas (Kozakiewicz et al. *in review*). The Interstate 5 (Figure 4.1) is one of the largest freeways in this region and has been shown to form a substantial barrier to bobcat gene flow but has also been shown to be permeable to FIV_{Lru} transmission (Lee et al., 2012; Fountain-Jones et al., 2017a). This discordance between host gene flow and FIV_{Lru} transmission indicates that rates of bobcat dispersal do not necessarily predict the likelihood of disease spread, reflecting greater sensitivity of the viral genetic structure to host movements (opportunities for viral gene flow exist additionally to host reproduction). Therefore, analysis of FIV_{Lru} may provide insights into both host and pathogen connectivity that are undetectable using host genetics or telemetry studies alone. These valuable studies have focused on a small region of southern California, on a single barrier, and a relatively small number of samples. Further work is required to test the generality of discordance between host connectivity and pathogen transmission, and to gain an understanding of factors that influence relationships among host and pathogen gene flow.

Here, we test whether patterns of pathogen genetic structure are consistent with known anthropogenic barriers to host gene flow. We sequence FIV_{Lru} isolated from fragmented bobcat populations across coastal Southern California (from north of Los Angeles to San Diego) to address this. Implementing Bayesian phylogenetics on the viral sequences, we quantify FIV_{Lru} connectivity and assess the permeability of several major highway and urban barriers to pathogen spread. In doing so, we aim to understand whether pathogen dynamics (as measured by gene flow) may differ across space and through time as urban development has expanded. We hypothesise that FIV_{Lru} transmission is occurring among bobcat populations in this region despite low host gene flow between habitat fragments. We

also expect that smaller and more isolated bobcat populations will have a lower diversity of FIV_{Lru} strains. This study provides novel insights into host and pathogen connectivity with respect to major urban barriers.

4.3 Methods

4.3.1 Study area, sample collection, and DNA extraction

Blood and tissue samples were collected from 292 bobcats in three adjoining areas in coastal Southern California (Figure 4.1). Forty-five were collected from north and east of San Diego between 2007–2012 (Jennings & Lewison, 2013), 113 were collected from southeast of Los Angeles between 2002–2010 (Lyren et al., 2006, 2008a, 2008b), and 134 were collected from northwest of Los Angeles between 1997–2011 (Riley et al., 2006; Serieys et al., 2015). Collectively, these samples represent five genetically distinct bobcat populations separated by major highways and urban areas that pose barriers to host gene flow (Kozakiewicz et al. *in review*). We label these populations according to their location with respect to major highway barriers: North-101, South-101, East-405, West-5, and East-5 (Figure 4.1). All samples were derived from either live trapping or opportunistically collected roadkill, and sample dates and locations recorded. Sex was recorded at the time of capture, and age (juvenile <2 years; adult >2 years) estimated according to size, weight, and dental wear. All live animals were captured, handled, and released according to protocols approved by cooperating agencies and relevant animal ethics authorities (for detailed information see Lyren et al., 2006; Riley et al., 2006; Lyren et al., 2008a, 2008b; Jennings & Lewison, 2013; Serieys et al., 2015). We extracted DNA using the Qiagen DNeasy Blood & Tissue Kit (Qiagen Inc., Valencia, CA, USA), eluting DNA in buffer EB.

4.3.2 FIV screening and amplification

Bobcats infected with FIV_{Lru} were identified by PCR assay of proviral DNA. We used nested PCR primers designed by Lee *et al*. (2014) to amplify 547 bp from a highly conserved region of the *gag* gene. First-round screening primers were A6F and A9R, and second-round screening primers were A7F and A8R (Lee et al., 2014). Screening PCRs were conducted in 50 µL reactions using Invitrogen Platinum Taq DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA), with first-round reactions containing 5 µL of DNA, which varied in concentration among samples between 10–100 ng/µL. Second-round reactions contained 2 µL of first-round PCR product. Reaction conditions included a hot start at 94°C for 2 min followed by 13 cycles comprising a melting temperature of 94°C for 30 s, touchdown annealing temperatures of 56°C–50°C (first round) or 59°C–53°C (second round) for 30 s and decreasing 0.5°C every cycle, and an extension temperature of 72°C for 30 s. These were followed by a further 27 cycles with the same conditions the previous cycle. A final extension temperature of 72°C was held for 5 min. PCRs were conducted in small batches of 16 to minimise risk of contamination, and each batch included two positive controls and one negative control. PCR product was visualised under UV light on a 1% agarose gel containing Gel Red (Gold Biotechnology Inc., Olivette, MO, USA). We identified a total of 73 FIV-positive bobcats, which we included in FIV_{Lru} genetic analysis below. Because the amplified *gag* gene region is highly conserved, we believe these to represent most of the sampled FIV-positive individuals, although some false negatives are likely (Kerr et al., 2018). Logistic regression with a likelihood ratio test (LRT) was used to explore trends in FIV_{Lru} prevalence among populations, sexes, and age categories.

For phylogeny reconstruction, we analysed a region of the highly variable *env* gene, which encodes the glycoproteins that constitute the retroviral envelope. PCR amplification of a 1.6 kbp fragment spanning a portion of the *env* gene to the 3’ long terminal repeat (LTR)

was conducted using nested PCR primers and protocols developed by Lee *et al.* (2014). First-round primers were A21F and A32R, which amplified a 2,267 bp fragment. Second-round primers were A22F and A31R, which amplified a 1,547 bp fragment. PCR was conducted in 50 μL reactions using Invitrogen Platinum Taq DNA Polymerase High Fidelity, with first-round reactions containing 5 μL of DNA as above, and second-round reactions containing 2 ng of first-round PCR product. First round reaction conditions included a hot start at 94°C for 1 min followed by 13 cycles comprising a melting temperature of 94°C for 30 s, touchdown annealing temperatures of 58°C–52°C for 30 s and decreasing 0.5°C every cycle, and an extension temperature of 68°C for 2.5 min. These were followed by a further 27 cycles with the same conditions except an annealing temperature of 52°C and then a final extension temperature of 68°C held for 3 min. Second round reaction conditions included a hot start at 94°C for 1 min followed by 17 cycles comprising a melting temperature of 94°C for 30 s, touchdown annealing temperatures of 59°C–51°C for 30 s and decreasing 0.5°C every cycle, and an extension temperature of 68°C for 1.5 min. These were followed by a further 23 cycles with the same conditions except an annealing temperature of 51°C and then a final extension temperature of 68°C held for 3 min.

Following second-round reactions and verification of amplification using gel electrophoresis as above, PCR product was purified using ExoSAP-IT PCR Product Cleanup Reagent (Thermo Fisher Scientific) according to the manufacturer's protocol. Purified PCR product was sequenced at Macrogen USA on an ABI 3730xl using internal forward and reverse primers (A23F and A23R; Lee et al., 2014) in addition to second round PCR primers above (A22F and A31R; Lee et al., 2014).

4.3.3 Sequence analysis and alignment

All visualisation and manipulation of sequence chromatograms were conducted using Geneious version 11.1.4. We manually screened all chromatograms to ensure they were correctly scored. We constructed a reference comprising a consensus of 12 existing Californian bobcat FIV_{Lru} sequences from Genbank (Lee et al., 2014) and mapped all reads to this, forming a new consensus sequence for each individual. Some FIV_{Lru} genetic variation within hosts was observed, indicated by multiple chromatogram peaks at the same base position across multiple reads. In such cases, the "dominant" strain was scored, identified as the highest quality peak at a given base position. If this method was unable to resolve the dominant genotype, the position was scored using the appropriate IUPAC degenerate character.

We used "Find ORFs" in Geneious to identify the *env* open reading frame (ORF) for each sequence. All sequences were trimmed to include only *env* to a final length of 1,257 bp, removing the sequenced 3' LTR. Stop codons were removed from the end of all ORFs and sequences were aligned using the MUSCLE translation alignment. The final alignment included sequences genotyped in this study together with 12 Genbank sequences trimmed from full genomes sequenced by Lee et al. (2014).

Aligned ORFs were examined for recombination breakpoints using RDP version 4.96 (Martin et al., 2015) with several recombination detection methods and assessing consensus among methods. Recombination detection methods used were RDP (Martin & Rybicki, 2000), GENECONV (Padidam et al., 1999), Chimaera (Posada & Crandall, 2001), MaxChi (Maynard Smith, 1992), BootScan (Salminen et al., 1995), SiScan (Gibbs et al., 2000), and 3Seq (Boni et al., 2007), with recombination breakpoints accepted if detected using more than two of these methods at a significance of $P < 0.05$. Any recombinant regions were

removed for subsequent analysis. To quantify genetic diversity among FIV_{Lru} strains within host populations, nucleotide diversity (π) was estimated using DnaSP (Rozas et al., 2017).

4.3.4 Phylogenetic analysis

A temporally explicit Bayesian phylogenetic tree was constructed using a 1,257 bp region of *env* with Bayesian Evolutionary Analysis Sampling Trees (BEAST) version 1.10, using BEAUti to construct the input files (Drummond et al., 2012). Tree tip dates were specified according to sampling date. We tested several models comprising different substitution models and molecular clocks and selected the most appropriate using marginal likelihood estimation with path and stepping-stone sampling (Baele et al., 2012). The highest supported model included the HKY substitution model with gamma distributed rate heterogeneity and a proportion of invariant sites, the two-partition codon partition model, and a lognormal uncorrelated relaxed molecular clock (see Appendix Table A4.1.1 for all model selection results). We ran three sets of 100 million Markov chain Monte Carlo iterations, sampling every 10,000 iterations and excluding the initial 10% of each set as burn-in. Model convergence was checked and parameters evaluated using Tracer version 1.7 (Rambaut et al., 2018). A maximum clade credibility tree was constructed from the sampled trees using TreeAnnotator version 1.10 from the BEAST package and visualised using FigTree version 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

In addition to reduced connectivity, phylogeographic structure can occur due to local adaptation, arising from spatial variation in host immunity or environmental conditions that can affect the fitness of viral lineages (Lion & Gandon, 2015). To identify signatures of positive selection acting across branches, we used the BUSTED (Murrell et al., 2015) method implemented in HyPhy (Pond & Muse, 2005), based on the maximum clade credibility tree

generated above. We tested all branches using aBSREL (Smith et al., 2015) implemented in HyPhy to identify specific branches upon which positive selection is acting, using a Holm-Bonferroni corrected *P*-value cut-off of 0.05.

We tested heterogeneity of FIV_{Lru} spread across highways using *geiger* (Harmon et al., 2008) in R (R Development Core Team, 2013). Treating host population as a discrete trait, we fitted three models that tested different hypotheses regarding relative rates of trait state change along the virus phylogeny: a) all transitions among states occur at the same rate (null model); b) all transitions among states occur at different rates; and c) all transitions among states occur at different rates but are symmetrical (e.g., transitions from state 1 to state 2 occur at the same rate as transitions from state 2 to state 1, but at a different rate from transitions involving other states). We conducted likelihood ratio tests of support for each variable-rate hypothesis in explaining the observed distribution of transitions among host populations across our virus maximum clade credibility tree, relative to the null model. These hypotheses were tested both before and after removal of branches undergoing selection to ensure that the spatial distribution of selection was not producing false signals of homogeneity (or heterogeneity) in road crossings.

4.4 Results

Total FIV_{Lru} prevalence across all sampled bobcats was 25.0% and did not vary significantly among the five host populations (21.7–27.3%; logistic regression LRT: G_1 = 0.72, P = 0.94). FIV_{Lru} infection was more likely among males than females (G_1 = 7.27, P < 0.01), and less likely among young individuals than in adults (G_1 = 7.83, P < 0.01), suggesting primarily horizontal transmission among males. Five individuals screened for

FIV_{Lru} lacked precise location data, including one FIV_{Lru}-positive individual, and were thus not included in the above estimates.

We obtained a total of 53 sequences for a 1,257 bp region of *env*, of which one lacked precise location data. We were unable to obtain sequences of sufficient quality for 20 FIV_{Lru}-positive samples, likely due to DNA degradation, low FIV_{Lru} copy number, or mutations at primer binding sites. Sequencing success did not appear to vary spatially or temporally. In addition, two sequences sampled from siblings at similar locations were identical, so only one was included in analyses for a total of 52. Recombination was not detected.

4.4.1 Phylogenetic analysis

Bayesian phylogenetic analysis indicated two major clades corresponding to FIV_{Lru} isolates from either northwest of Los Angeles (North-101, South-101, and East-405 populations), or southeast of Los Angeles (East-5 and West-5 populations; Figure 4.2; Table 4.1). These two lineages diverged from a common ancestor in approximately the year 1901 (95% highest posterior density; HPD = 1810–1967), and this node had high posterior support (*PP* = 1.00). The clade from southeast of Los Angeles was characterised by relatively deeper divergences among isolates, with many long branches comprising relatively few isolates each. The two longest of these branches diverged prior to 1925 but had relatively low posterior support. The two largest subclades southeast of Los Angeles diverged around 1929 (95% HPD = 1863–1975) with relatively high posterior support (*PP* = 0.98).

Within the clade southeast of Los Angeles, there was an absence of phylogeographic structure in FIV_{Lru} with respect to host population boundaries or other geographic features, although geographically proximal isolates were sometimes closely related (Figure 4.1b, Figure 4.2). Only two FIV_{Lru} isolates from the West-5 population were more closely related

to each other than to isolates from East-5. These West-5 isolates diverged from East-5 isolates in approximately 1983 (95% HPD = 1962–1999). Two further divergences among West-5 and East-5 isolates occurred in approximately 1932 (95% HPD = 1851–1987) and 1996 (95% HPD = 1988–2001).

In contrast to the clade southeast of Los Angeles, FIV_{Lru} isolates sampled northwest of Los Angeles show strong phylogeographic structure. This clade forms two major groups that diverged, with high posterior support (*PP* = 1.00), in approximately 1950 (95% HPD = 1898–1986). This divergence is among groups comprising all isolates sampled from the South-101 and East-405 populations, and another group containing all FIV_{Lru} isolates sampled from the North-101 population. Three isolates sampled from southeast of Los Angeles diverged from the northwest Los Angeles group more recently than others – in approximately 1923 – but with only moderate posterior support (*PP* = 0.68). These three isolates (two from East-5 and two from West-5) were geographically dispersed and not proximal to the northwest of Los Angeles populations.

4.4.2 Sequence diversity, selection, and highway crossing rate variation

FIV_{Lru} sequences were most diverse southeast of Los Angeles (Figure 4.3), with East-5 (π = 0.095; SD = 0.005) and West-5 (π = 0.084; SD = 0.022) having higher nucleotide diversity than North-101 (π = 0.014; SD = 0.002), South-101 (π = 0.026; SD = 0.003), and East-405 (π = 0.008; SD = 0.001). BUSTED analysis showed evidence for positive selection acting on our phylogeny ($P < 0.0001$). aBSREL tested all 103 branches for evidence of positive selection and identified two branches showing signatures of selection (Figure 4.2). One of these branches contained only two FIV_{Lru} isolates from the South-101 population that was sister to all other isolates sampled from that population. The other branch showing

evidence for positive selection was the large branch comprising a majority of the FIV_{Lru} sampled from southeast of Los Angeles that diverged into two large subclades in approximately 1929.

Likelihood ratio tests of trait evolutionary hypotheses (i.e., testing for variation in rates of highway crossing while accounting for variation in phylogenetic branch depth) indicated no support for either the variable symmetrical rate ($P > 0.05$, $x^2 = 12.4$, $df = 19$) or variable asymmetrical rate ($P > 0.05$, $x^2 = 8.32$, $df = 9$) state change models over the equal rate state change model. Tests following removal of branches undergoing selection produced similar results. Therefore, the observed patterns of FIV_{Lru} spread across major highways are best explained by equal rates of FIV_{Lru} exchange among each pair of host populations.

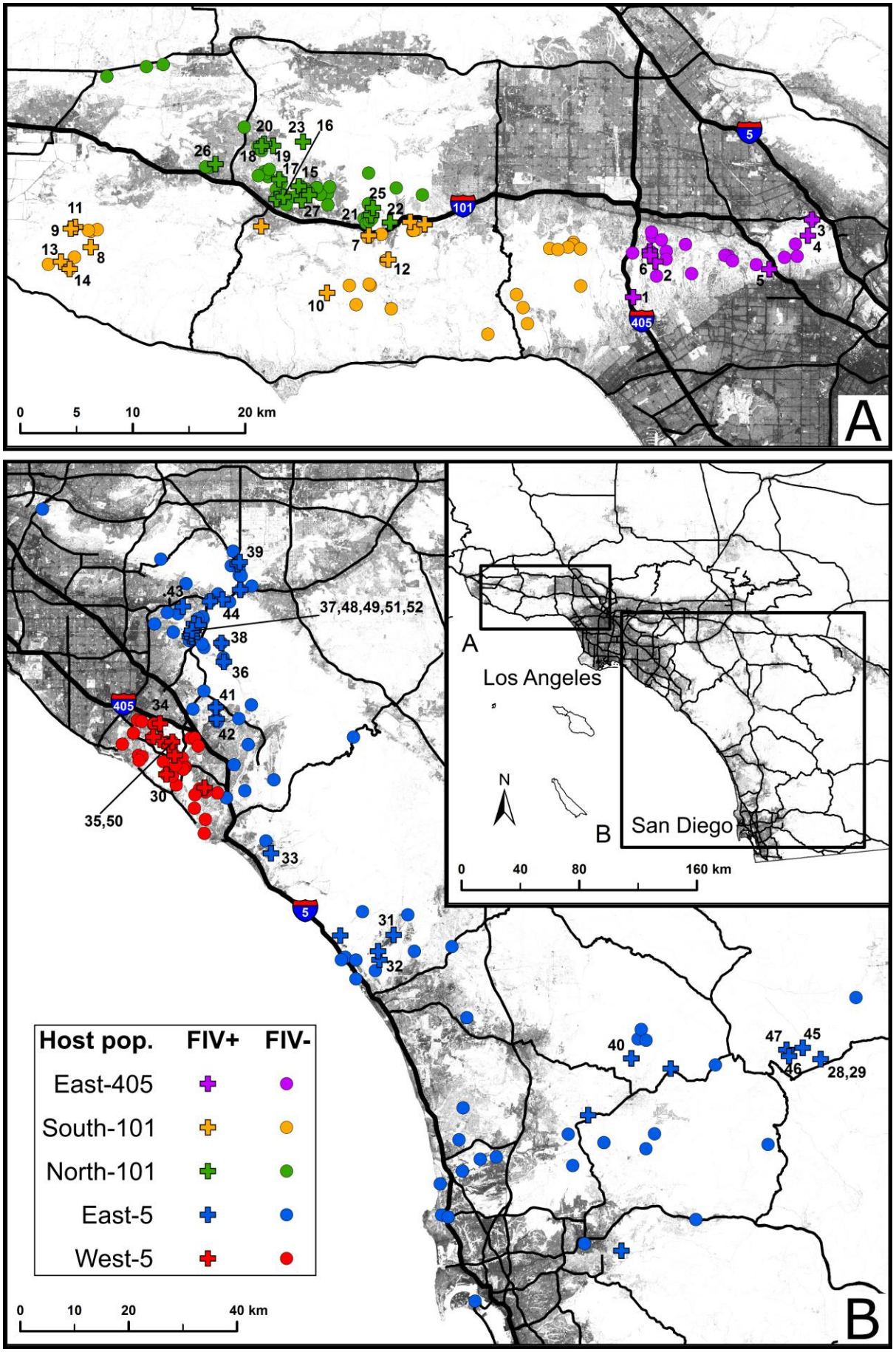
A
0 5 10 20 km
B
Los Angeles
San Diego
A
B
N
0 40 80 160 km
0 10 20 40 km
Host pop. FIV+ FIV-
East-405
South-101
North-101
East-5
West-5
37,48,49,51,52
35,50
28,29
405
5
101

Figure 4.1. Bobcat sample locations in coastal southern California: (A) northwest of Los Angeles, and (B) southeast of Los Angeles. Grey shading indicates urban development, including the cities of Los Angeles and San Diego (see inset). Black lines indicate highways, with highways acting as barriers to gene flow among host populations labelled with blue-and-red route markers and indicated as bold black lines. Marker colours indicate host population, with crosses indicating FIV-positive individuals and circles indicating FIV-negative individuals. All FIV-positive samples included in phylogenetic analysis are numbered for cross-referencing with Figure 4.2. FIV-positive samples that lack numbers indicate unsuccessful genotyping.

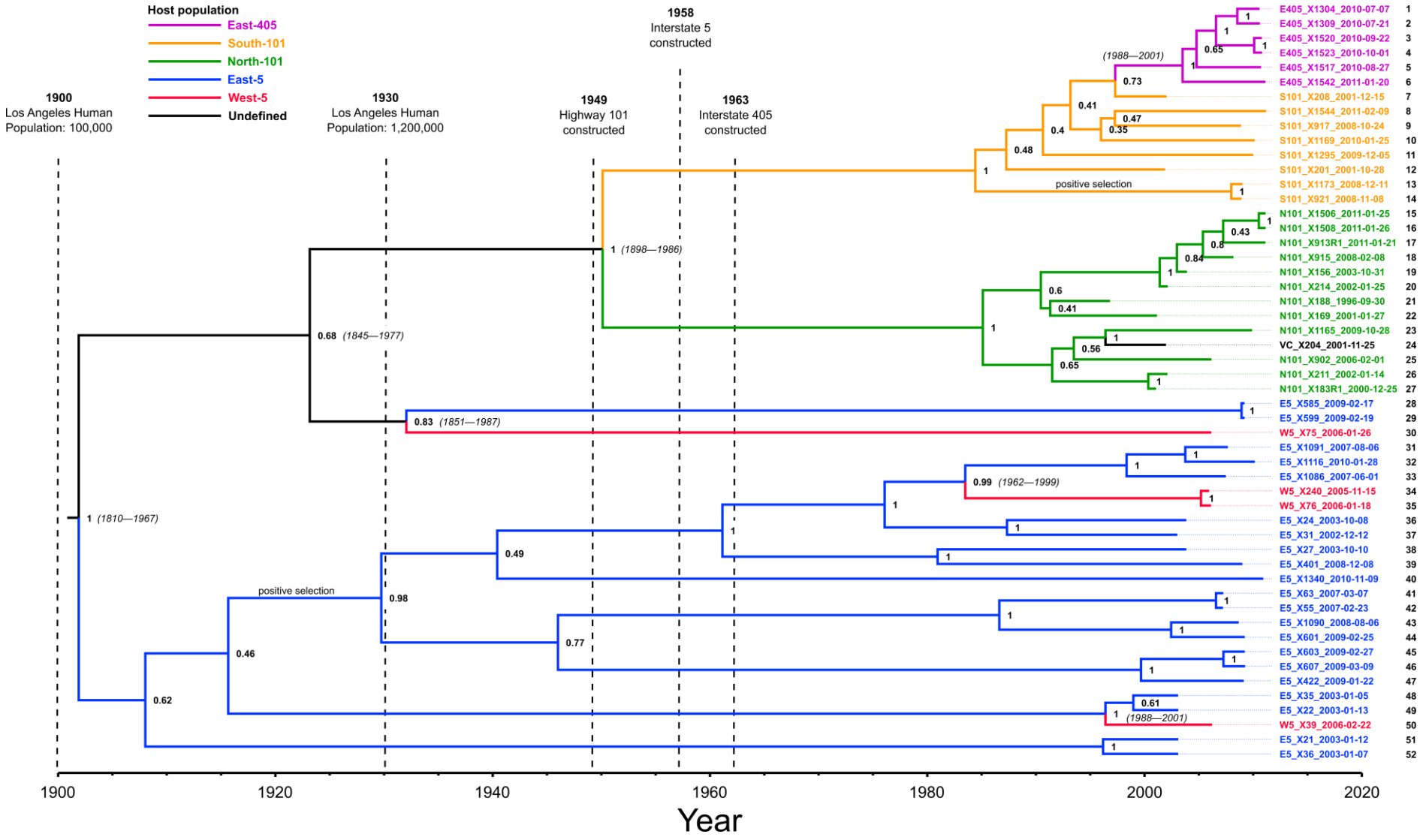
Host population
East-405
South-101
North-101
East-5
West-5
Undefined
1900
Los Angeles Human
Population: 100,000
1930
Los Angeles Human
Population: 1,200,000
1949
Highway 101
constructed
1958
Interstate 5
constructed
1963
Interstate 405
constructed
1 (1810—1967)
0.68 (1845—1977)
1 (1898—1986)
0.83 (1851—1987)
positive selection
0.46
0.62
0.98
0.49
0.77
0.99 (1962—1999)
(1988—2001)
0.61
(1988—2001)
positive selection
E405_X1304_2010-07-07 1
E405_X1309_2010-07-21 2
E405_X1520_2010-09-22 3
E405_X1523_2010-10-01 4
E405_X1517_2010-08-27 5
E405_X1542_2011-01-20 6
S101_X208_2001-12-15 7
S101_X1544_2011-02-09 8
S101_X917_2008-10-24 9
S101_X1169_2010-01-25 10
S101_X1295_2009-12-05 11
S101_X201_2001-10-28 12
S101_X1173_2008-12-11 13
S101_X921_2008-11-08 14
N101_X1506_2011-01-25 15
N101_X1508_2011-01-26 16
N101_X913R1_2011-01-21 17
N101_X915_2008-02-08 18
N101_X156_2003-10-31 19
N101_X214_2002-01-25 20
N101_X188_1996-09-30 21
N101_X169_2001-01-27 22
N101_X1165_2009-10-28 23
VC_X204_2001-11-25 24
N101_X902_2006-02-01 25
N101_X211_2002-01-14 26
N101_X183R1_2000-12-25 27
E5_X585_2009-02-17 28
E5_X599_2009-02-19 29
W5_X75_2006-01-26 30
E5_X1091_2007-08-06 31
E5_X1116_2010-01-28 32
E5_X1086_2007-06-01 33
W5_X240_2005-11-15 34
W5_X76_2006-01-18 35
E5_X24_2003-10-08 36
E5_X31_2002-12-12 37
E5_X27_2003-10-10 38
E5_X401_2008-12-08 39
E5_X1340_2010-11-09 40
E5_X63_2007-03-07 41
E5_X55_2007-02-23 42
E5_X1090_2008-08-06 43
E5_X601_2009-02-25 44
E5_X603_2009-02-27 45
E5_X607_2009-03-09 46
E5_X422_2009-01-22 47
E5_X35_2003-01-05 48
E5_X22_2003-01-13 49
W5_X39_2006-02-22 50
E5_X21_2003-01-12 51
E5_X36_2003-01-07 52
1900
1920
1940
1960
1980
2000
2020
Year

Figure 4.2. Maximum clade credibility Bayesian phylogenetic tree constructed using 1257 bp sequences of the env gene region in FIV sampled from bobcats in coastal southern California, with dated tips. Node labels indicate posterior probabilities, with 95% highest probability density estimates for divergence dates shown in parentheses for nodes associated with among-population transmission. Branches inferred to be undergoing positive selection are labelled as such. Branch and tip label colours and tip label prefixes indicate host population, and dates of major urban features corresponding to host population structure are shown. All tips are numbered for cross-referencing with sample locations in Figure 4.1.

Table 4.1. Among-host-population phylogeographic divergences of bobcat feline immunodeficiency virus and potentially associated urban influences.

Year	Divergence event
1901	• FIV strains from northwest of Los Angeles diverged from strains southwest of Los Angeles. Los Angeles human population approx. 100,000.
1923	• An additional lineage containing three contemporary isolates sampled southeast of Los Angeles diverged from the group northwest of Los Angeles. Los Angeles human population approx. 1,000,000.
1932	• First West-5 lineage diverged from East-5 population (prior to construction of Interstate 5).
1950	• Divergence among lineages sampled north of highway 101 and those sampled south of 101. Construction of the 101 had been completed a year earlier in 1949.
1983	• Second West-5 lineage diverged from East-5 population, 25 years after the Interstate 5 was constructed in 1958.
1996	• Third West-5 lineage diverged from East-5 population.
1997	• All contemporary isolates sampled east of the Interstate 405 diverged from those sampled west of Interstate 405, 34 years after the 405 was constructed in 1963.

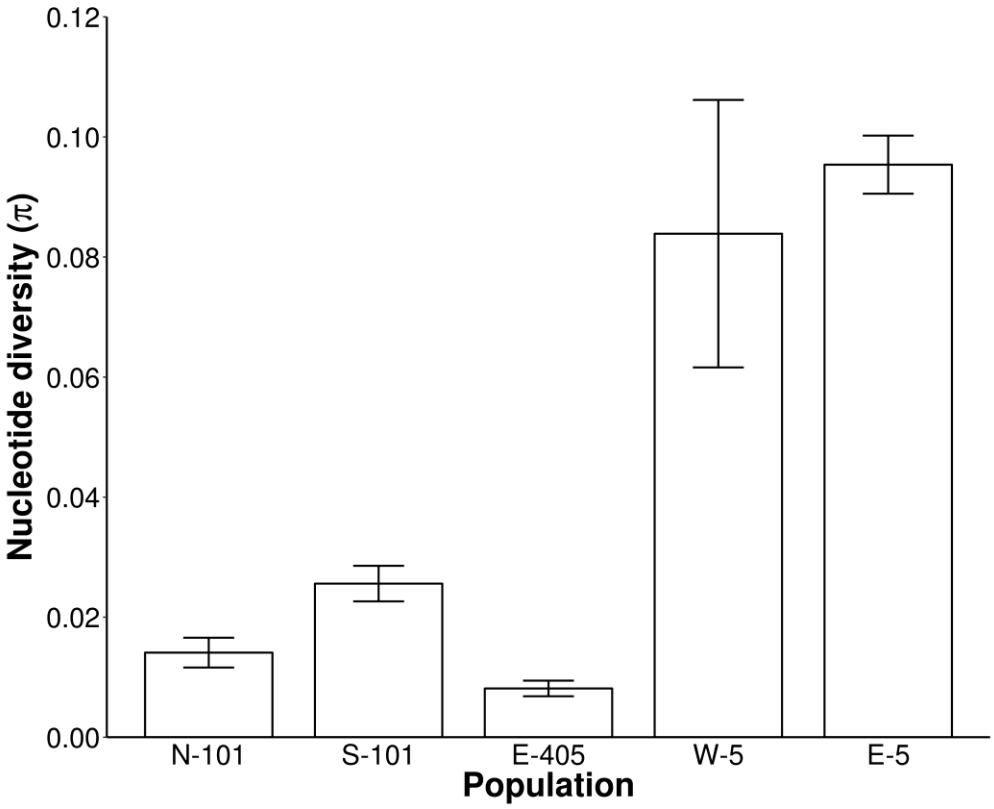

Figure 4.3. FIV nucleotide diversity (±SD) is higher in bobcat populations located southeast of Los Angeles (West-5 and East-5) than those located northwest of Los Angeles (North-101, South-101, and East-405).

4.5 Discussion

Understanding how pathogens may spread among host populations putatively isolated by anthropogenic barriers is important for assessing the vulnerability of regional populations to future disease outbreaks. To elucidate these processes, we conducted one of the most spatially intensive viral phylogeographic studies in an urban environment, enabling us to assess phylogeographic patterns among several populations. We found distinct differences in FIV_{Lru} phylogenetic structure among populations in different regions in coastal southern California. In general, populations southeast of Los Angeles had greater divergence among FIV_{Lru} lineages, high FIV_{Lru} genetic diversity, and lacked clear phylogeographic structure. In contrast, the populations northwest of Los Angeles are relatively more isolated, with all FIV_{Lru} isolates sampled north of highway 101 monophyletic and those sampled south of highway 101 paraphyletic for FIV_{Lru} isolated from the South-101 and East-405 populations. This indicates that for each host population northwest of Los Angeles, all FIV_{Lru} isolates within a given host population descended from a single founding isolate.

The observed phylogenetic relationships indicate that major highways are limiting FIV_{Lru} connectivity among host populations around Los Angeles. Despite marked differences in phylogenetic structure northwest and southeast of Los Angeles, we found no evidence of any difference in rates of FIV_{Lru} movement across highways among each of these areas. Instead, the greater number of observed highway crossings southeast of Los Angeles is likely due to the greater depth of branches here compared to northwest of Los Angeles, thus having a greater probability of observing crossings over this time. Compared to populations northwest of Los Angeles, the East-5 FIV_{Lru} population has substantially higher genetic diversity as well as a high population size owing to a large host population size and its openness to unsampled populations to the east that may serve as potential sources for diverse FIV_{Lru} strains (Kozakiewicz et al. *in review*). Our results suggest that much of this genetic

diversity was retained in the West-5 FIV_{Lru} population when its host population became isolated.

Incomplete lineage sorting occurs when the phylogenetic tree derived from the genotyping of single strains differs from the population-level transmission tree. This is common in viral phylodynamics, and is more likely to occur when intra-host viral diversity is high (Volz et al., 2017), which is plausible southeast of Los Angeles due to the high diversity of FIV_{Lru} strains overall. In addition, a large effective population size is a key factor contributing to incomplete lineage sorting (Pamilo & Nei, 1988; Maddison, 1997; Pease & Hahn, 2013). Thus, incomplete lineage sorting may explain the observed lack of phylogeographic structure among East-5 and West-5 FIV_{Lru} isolates. Indeed, incomplete lineage sorting has been suggested to have caused incongruence among host and parasite pathogen genetic structure in other systems (Symula et al., 2011; Torres-Pérez et al., 2011).

Because FIVs are directly transmitted, movements of FIV strains reflect the movements of infected hosts. Thus, by phylogenetically estimating divergence dates among virus strains, we can infer how historical patterns of urbanisation have potentially influenced bobcat connectivity over time. Estimated divergence dates with respect to major anthropogenic barriers generally reflected the chronology of urban development in coastal southern California. The divergence of the two major FIV_{Lru} clades estimated at around 1901 coincides roughly with Los Angeles' initial development boom when its human population expanded from approximately 6,000 to 100,000 between 1870–1900 (U.S. Census Bureau, 1900). It is unlikely that at this time the extent of urban development in the Los Angeles Basin was alone sufficient to preclude movement of bobcats among the areas inhabited by the two clades we identified. However, a large proportion of anthropogenic land use associated with Los Angeles at this time consisted of agriculture (Nelson, 1959), which would also have constituted removal of large areas of bobcat habitat and has been suggested to reduce

connectivity in contemporary populations (Hilty & Merenlender, 2004). Our phylogenetic analysis provides evidence for minimal transmission across this region since ca. 1901, with only one more recent divergence evident among FIV_{Lru} populations northwest and southeast of Los Angeles (1923).

Highway 101 is a primary driver of host population structure in bobcats northwest of Los Angeles (Riley et al., 2006; Serieys et al., 2015), and its completion in 1949 corresponds to a divergence in FIV_{Lru} estimated to have occurred in approximately 1950. Highway 101 has thus constituted a major barrier to FIV_{Lru} transmission since its construction, with no evidence in contemporary populations of any FIV_{Lru} strains having crossed this road. In contrast, previous research has indicated some highway crossing by bobcats, particularly from south-to-north, evidenced from both genetic population assignment as well as telemetry (Riley et al., 2006), with subsequent support from genetic studies (Kozakiewicz et al. *in review*; Serieys et al., 2015). In theory, these movements should facilitate FIV_{Lru} transmission, which occurs predominantly through aggressive or sexual contact (VandeWoude & Apetrei, 2006). However, coalescent simulation has suggested bobcats that cross highway 101 rarely reproduce (Riley et al., 2006), potentially indicating that immigrants are infrequently capable of sexually transmitting FIV_{Lru}, which may ultimately lead to reduced transmission among populations overall.

A group of six isolates sampled from the East-405 population diverged from the majority of isolates in the South-101 clade in approximately 1997, providing evidence that the Interstate 405 is a barrier to FIV_{Lru} transmission in addition to host gene flow (Kozakiewicz et al. *in review*; Serieys et al., 2015). The Interstate 405 was completed in 1962, and the 95% HPD intervals for this divergence estimate only extends to 1988. It is therefore probable that transmission occurred among the South-101 and East-405 populations for some time following the completion of the Interstate 405, with bobcats potentially

utilising underpasses where I-405 crosses over minor roads. Host genetic studies also provide evidence of some movement across I-405 (Kozakiewicz et al. *in review*; Serieys et al., 2015), providing further indication of potential for FIV_{Lru} spread, though apparently of insufficient frequency for multiple FIV_{Lru} lineages of South-101 origin to establish in East-405.

The other major barrier known to be driving bobcat host population structure is the Interstate 5 (Kozakiewicz et al. *in review*; Lee et al., 2012; Ruell et al., 2012), located southeast of Los Angeles. Our results suggest that despite little visible phylogeographic structure with respect to these host populations, rates of FIV_{Lru} exchange are low and comparable to those among populations northwest of Los Angeles. However, we did observe three lineages of FIV_{Lru} in the West-5 population that originated in the East-5 population, diverging in approximately 1932, 1983, and 1996, consistent with previous studies in this area (Lee et al., 2012; Fountain-Jones et al., 2017a). The latter two of these events occurred following the completion of Interstate 5 in 1958, suggesting that some transmission (and thus bobcat migration) across this highway has occurred since its completion. Several other large highways run through our study area southeast of Los Angeles, including the Interstate 15, which constitutes a barrier to gene flow in mountain lions (Ernest et al., 2014) but not bobcats (Kozakiewicz et al. *in review*). We observed no FIV_{Lru} phylogeographic structure with respect to these highways.

Spatial heterogeneity in either environmental factors or host immunology may create variation in conditions to which pathogens must adapt, and may influence pathogen phylogeographic patterns, due to, for example, higher fitness of locally adapted strains resulting in increased pathogen population structure (Kaltz & Shykoff, 1998). The gene region under study, *env*, encodes the glycoproteins that constitute the retroviral envelope, which serves a protective function and enables incorporation into host cells. Thus, it is thought to be under more immunological selection pressure compared to other parts of the

FIV genome, with some amino acid sites evolving under positive selection (Lee et al., 2014). We explored the potential of local adaptation to explain the observed phylogeographic patterns and found no evidence of positive selection acting on any branches associated with the highly structured FIV_{Lru} populations northwest of Los Angeles. However, we did find evidence for positive selection acting on two lineages, including a branch that diverged into two large clades comprising a majority of FIV_{Lru} isolates from southeast of Los Angeles. These clades do not reflect any clear spatial pattern but do encompass a large geographical area containing a large, genetically diverse host population (Kozakiewicz et al. *in review*). Although it is unclear what factors are driving this signal of selection, variation in host immunity across the southeast area may be generating selective pressure on FIV_{Lru} lineages circulating here.

In conclusion, our results demonstrate that the degree of congruence among host and pathogen spatial population structure can vary among populations with comparable landscape characteristics. However, we show that this incongruence may be due to inherent factors such as incomplete lineage sorting occurring where viral populations are larger and more diverse, rather than necessarily variation in pathogen connectivity. In general, FIV_{Lru} rarely crosses major barriers to host connectivity, and this rate does not vary regionally. This infrequent crossing indicates that these populations should be managed as separate units in the event of a serious disease outbreak. Our results suggest that direct assessment of pathogen transmission is important for estimating the ability of geographic barriers to constrain disease outbreaks, but that pathogen phylogenetic structure (or lack thereof) should be carefully scrutinised before conclusions are drawn about rates of transmission among contemporary host populations.

4.6 Appendix 1

Table A4.1.1 BEAST model selection results for tests of substitution models and molecular clocks. Log-maximum likelihood scores for path sampling and stepping-stone sampling are shown.

Substitution model	Molecular clock	Log ML (PS)	Log ML (SS)
GTR	strict	-8457.96	-8457.72
GTR	uncorrelated relaxed	-8354.26	-8354.1
HKY	strict	-8477.64	-8477.39
HKY	uncorrelated relaxed	-8351.76	-8352.04

4.7 Appendix 2

To further characterise FIV_{Lru} sequence diversity, we used Sequence Demarcation Tool v1.2 (SDT; Muhire et al., 2014) to identify genetic clusters among sequences. With SDT, we aligned sequences using MUSCLE and generated neighbour-joining trees, from which matrices of pairwise sequence identity were produced. Separately for each of the northwest and southeast of Los Angeles groups, we calculated the mean pairwise percentage identity and used this as a threshold to define genetic clusters, with sequences having a pairwise identity above this threshold considered a distinct genetic cluster. By assessing the number of genetic clusters and the number of sequences obtained for each, we qualitatively estimated the degree to which our sampling captured the overall genetic diversity in each region.

For the 27 sequences sampled northwest of Los Angeles, the mean pairwise identity was 94% (Figure A4.2.1A). Using this threshold, we identified two genetic clusters (Figure A4.2.1B). These clusters segregated clearly according to their location either side of highway 101, with all 13 sequences sampled north of 101 forming a distinct group, and the 14 sampled south of 101 (including those from east of Interstate 405) forming another.

For the 25 sequences sampled southeast of Los Angeles, the mean pairwise identity was 90% (Figure A4.2.1C). Using this threshold, we identified seven genetic clusters (Figure A4.2.1D). One cluster comprised approximately half (12) of the sequences, with a second cluster comprising six sequences. However, the remaining seven sequences formed five clusters comprising only one or two sequences each.

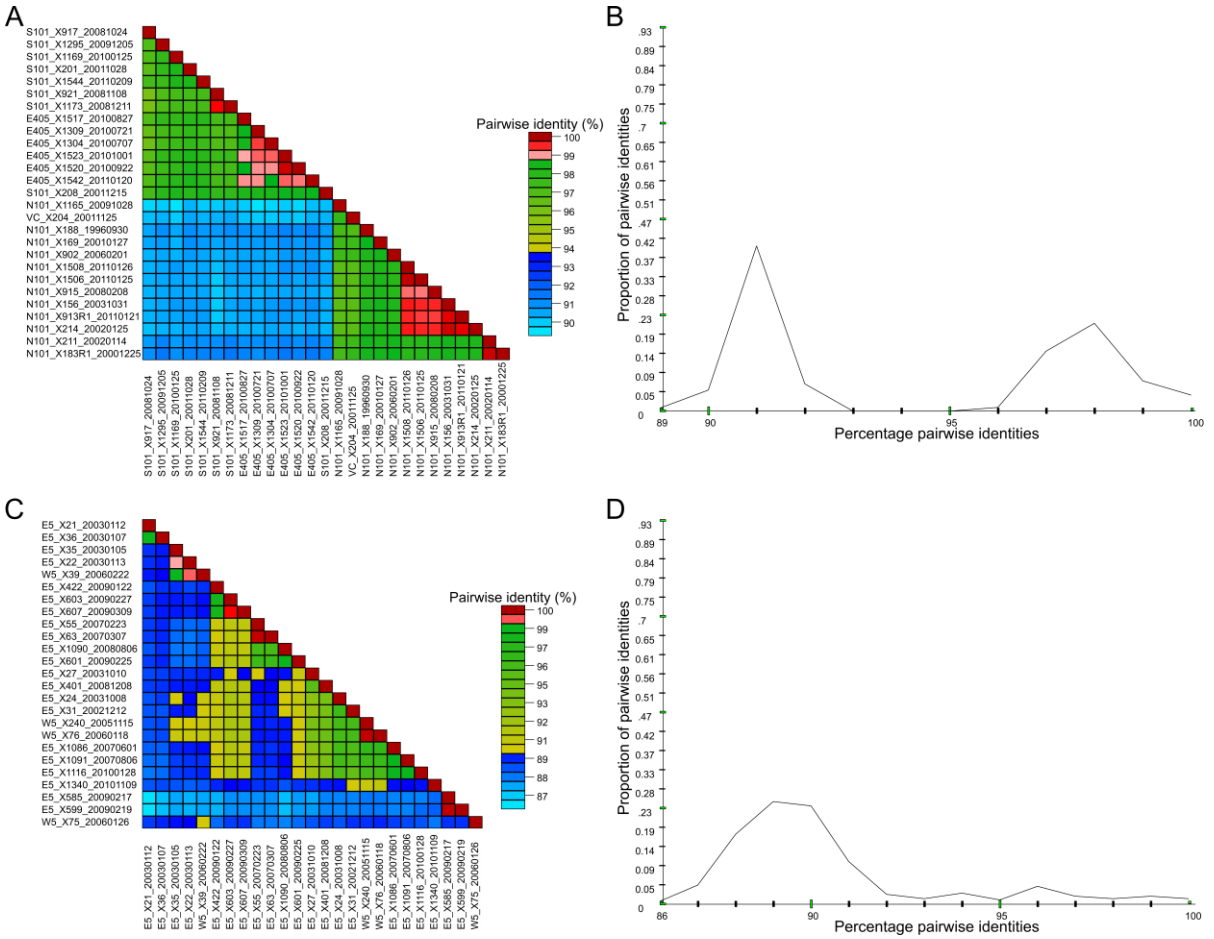

Figure A4.2.1 Pairwise identity among FIV$_{Lru}$ sequences northwest (A & B) and southeast (C & D) of Los Angeles. Higher pairwise identity (mean = 94%) was observed northwest of Los Angeles with two distinct sequence clusters, compared to a mean pairwise identity of 90% and seven sequence clusters southeast of Los Angeles.

Northwest of Los Angeles, there was a higher degree of pairwise identity overall and both genetic clusters identified here were well-represented in the data. The absence of any highly divergent individual samples suggests that our sampling is likely representative of the overall genetic variation northwest of Los Angeles. In contrast, southeast of Los Angeles, the five highly divergent individual or pairs of sequences may represent FIV$_{Lru}$ variants that are present in this landscape but are poorly represented in the data. The presence of poorly-represented clusters would also indicate a likelihood that there are additional clusters that were not sampled at all. However, bobcat populations southeast of Los Angeles are relatively

open, with no obvious geographical barriers preventing immigration from unsampled populations to the east. This immigration may infrequently bring novel strains of FIV_{Lru} into the sampled area, so it is possible that the observed low-sample clusters represent such events and thus reflect low prevalence of these particular strains within the sampled area. Because we do not have samples from other potential source populations, we cannot properly assess the provenance of these highly divergent sequences.

Chapter 5

Virus phylogeography in urban bobcat populations reveals increased transmission in vegetated habitats

5.1 Abstract

Spatially heterogeneous landscape factors such as urbanisation can have substantial impacts on the severity and spread of wildlife diseases. It is important to understand how these impacts may vary among populations and among different spatial scales, but research linking patterns of transmission to heterogeneous landscapes remains rare. We test the influence of landscape heterogeneity and host factors on phylogeographic estimates of feline immunodeficiency virus transmission (FIV_{Lru}) among bobcats (*Lynx rufus*) sampled from coastal southern California. We investigated local patterns of FIV_{Lru} transmission in populations northwest and southeast of Los Angeles, as well as broader, regional-scale patterns. We consistently found that natural habitat is an important factor shaping FIV_{Lru} transmission dynamics. Specifically, bobcats sampled from areas that differed in the amounts of natural land cover (i.e., forest or scrub) had more distantly related FIV_{Lru} strains. Because natural land cover was negatively associated with urban development, we suggest that the observed patterns may reflect reduced dispersal of bobcats from natural to more urbanised areas resulting in reduced FIV_{Lru} exchange among these areas. We also evaluated dispersal velocities of FIV_{Lru} lineages and found a positive association with vegetation density at the regional scale, further supporting the importance of natural habit in FIV_{Lru} transmission. These multiple lines of evidence demonstrate how urbanisation can change patterns of disease transmission and provide insights into how continued urban development may influence the incidence and management of wildlife disease.

5.2 Introduction

Understanding how host and landscape factors influence disease transmission is a focus of epidemiology and disease ecology (Gottdenker et al., 2014; Becker et al., 2015). Because landscape features (including anthropogenic landscapes) are spatially heterogeneous, we see spatial heterogeneity in pathogen transmission networks (i.e., the distribution of transmission events among individuals or populations) that can influence the severity and spread of disease outbreaks (Ostfeld et al., 2005; Meentemeyer et al., 2012). For example, the impacts of urbanisation on pathogen dynamics can be non-uniform, varying depending on the populations and the spatial scale under study (Becker et al., 2015; Bourhy et al., 2016; Dellicour et al., 2017). However, research linking pathogen transmission to heterogeneous landscape factors remains rare, particularly across replicated landscapes, but is necessary to understand how spatial patterns of disease are modified.

Directly relating pathogen transmission to specific landscape factors has been challenging because observing transmission events in wild populations is essentially impossible. Inference of contacts among hosts (often by measuring spatial overlap) has been a valuable proxy for transmission, but often requires intensive effort to conduct and may not accurately reflect true transmission events (Craft, 2015; Gilbertson et al., 2018). However, modern advances in molecular data collection and analysis are enabling researchers to quantify past transmission more precisely (Archie et al., 2009; Biek & Real, 2010; Didelot et al., 2017; Kozakiewicz et al., 2018a). This improves our ability to elucidate the impacts of landscape features on pathogen dynamics at multiple spatial scales within and among populations.

Viruses can be excellent candidates for the genetic inference of transmission. Their small genomes and rapid evolutionary rates produce generally high degrees of genetic

variation across transmission networks over ecologically relevant timescales. This has enabled researchers to understand where and when pathogen outbreaks have spread in the recent past, identifying transmission among extant populations or even specific individual hosts (Archie et al., 2009; Brunker et al., 2012; Biek et al., 2015), leading to advancements in our understanding of the transmission in a number of important pathogens; for example, MERS (Corman et al., 2014), Ebola (Carroll et al., 2015), influenza (Magee et al., 2015), HIV (Faria et al., 2014), and rabies (Streicker et al., 2016). However, it is only recently that tools have been developed to quantify how specific heterogeneous landscape factors shape viral transmission at fine spatiotemporal scales. These 'eco-phylogenetic' approaches integrate phylogenetics, phylodynamics (the study of processes shaping viral phylogenies; Grenfell et al., 2004), and community ecological methods to understand how host and pathogen communities and environmental factors interact to shape pathogen dynamics (Fountain-Jones et al., 2017c).

In addition to genetic variability, several other factors contribute to the utility of a given virus as a model for studying disease spread in wildlife populations (Kozakiewicz et al., 2018a). For example, viruses that are chronically infectious and either have low-pathogenicity or are apathogenic (and thus having little influence on host survival) enable populations to be sampled at any time post-infection. Further, species-specificity results in transmission networks that are largely contained within the sampled population (i.e., no transmission by other species). Viruses exhibiting such characteristics are well suited to target specific ecological questions of broad relevance to pathogen dynamics in particular hosts or environments (Kozakiewicz et al., 2018a). An example of such a virus is feline immunodeficiency virus (FIV), a fast-evolving RNA retrovirus that forms species-specific infections in many wild felids (VandeWoude & Apetrei, 2006; Carver et al., 2016). FIV has been relatively well studied, and phylogenetic analysis of FIV has provided insights into host

and pathogen dynamics in several species (Biek et al., 2006; Lee et al., 2012; Fountain-Jones et al., 2017a, 2017b).

FIV is endemic in many bobcat (*Lynx rufus*) populations (species-specific strain; FIV_{Lru}) and is thus well-suited to inferring impacts of host and landscape factors on pathogen transmission. Furthermore, bobcats are useful indicators of connectivity in areas where urban habitat fragmentation is occurring (Kozakiewicz et al. *in review*; Crooks, 2002). In coastal southern California (one of the most highly urbanised regions in North America), fragmented tracts of natural habitat support isolated bobcat populations. Major highways and dense tracts of urban development form the primary barriers to connectivity among populations (Kozakiewicz et al. *in review*; Riley et al., 2006; Lee et al., 2012; Ruell et al., 2012; Serieys et al., 2015). Although bobcats tend to avoid highly urban areas, they can tolerate a variety of habitats and degrees of anthropogenic disturbance (Riley et al., 2003; Lyren et al., 2008b), enabling them to persist in urban coastal southern California.

Here, we test whether host and fine-scale landscape factors influence pathogen transmission and spread, using FIV_{Lru} in coastal southern California as a model. In contrast to previous work in this region (Lee et al., 2012; Fountain-Jones et al., 2017a), we aim to elucidate how selected factors influencing pathogen transmission and spread can differ in local areas compared to broader, region-wide patterns (i.e., among fine and broad spatial scales), and among host populations experiencing different urban impacts. We focus on two components of FIV_{Lru} transmission. Firstly, we assess how host factors (including sex and relatedness) and urban and non-urban landscape factors (both at and among host sample locations) influence past rates of transmission, using FIV_{Lru} phylogenetic relatedness as a proxy. Second, we quantify rates and routes of pathogen spread across the landscape and identify landscape factors affecting how quickly FIV_{Lru} lineages disperse. These complementary approaches provide detailed insights into how landscape factors shape

spatiotemporal patterns of pathogen transmission in one of the most urbanised regions in North America.

5.3 Methods

We collected blood and tissue samples from 292 bobcats from three adjoining areas in coastal Southern California. Forty-five were collected from north and east of San Diego between 2007–2012 (Jennings & Lewison, 2013), 113 were collected from southeast of Los Angeles (LA) between 2002–2010 (Lyren et al., 2006, 2008a, 2008b), and 134 were collected from northwest of LA between 1996–2011 (Riley et al., 2006; Serieys et al., 2015). Collectively, these samples represent five genetically distinct bobcat populations separated by major highways and urban areas that pose barriers to host gene flow (Kozakiewicz et al. *in review*). We label these populations according to their location with respect to the major highway barriers: North-101, South-101, East-405, West-5, and East-5 (Figure 5.1). All samples were derived from either live trapping or opportunistically collected roadkill. For each sample, we recorded date, location, and sex. All live animals were captured, handled, and released according to protocols approved by cooperating agencies and relevant animal ethics authorities (for detailed information see Lyren et al., 2006; Riley et al., 2006; Lyren et al., 2008a, 2008b; Jennings & Lewison, 2013; Serieys et al., 2015). We extracted DNA using the Qiagen DNeasy Blood & Tissue Kit (Qiagen Inc., Valencia, CA, USA) according the manufacturers protocols but eluting DNA in buffer EB.

5.3.1 FIV_{Lru} screening and amplification

Bobcats infected with FIV_{Lru} were identified by PCR assay of proviral DNA according to Kozakiewicz *et al*. (*in prep*), using nested PCR primers designed by Lee *et al*.

(2014) to amplify 547 bp from a highly conserved region of the *gag* gene. For phylogenetic reconstruction, we analysed a region of the highly variable *env* gene, which encodes the glycoproteins that constitute the retroviral envelope. PCR amplification of a 1.6 kbp fragment spanning a portion of the *env* gene to the 3' long terminal repeat (LTR) was conducted using nested PCR primers and protocols developed by Lee *et al.* (2014) and described in Kozakiewicz *et al*. (*in prep*). PCR product was purified using ExoSAP-IT PCR Product Cleanup Reagent (Thermo Fisher Scientific) according to the manufacturer's protocol. Purified PCR product was sequenced at Macrogen USA on an ABI 3730xl using internal forward and reverse primers as described in Kozakiewicz *et al*. (*in prep*).

5.3.2 Sequence analysis and alignment

Sequences were trimmed to include only the *env* open reading frame (ORF), excluding stop codons, for a final length of 1,257 bp. The final alignment included 51 FIV_{Lru} sequences. Aligned ORFs were examined for recombination breakpoints using RDP version 4.96 (Martin et al., 2015) with several recombination detection methods: RDP (Martin & Rybicki, 2000), GENECONV (Padidam et al., 1999), Chimaera (Posada & Crandall, 2001), MaxChi (Maynard Smith, 1992), BootScan (Salminen et al., 1995), SiScan (Gibbs et al., 2000), 3Seq (Boni et al., 2007). Recombination breakpoints were accepted if detected using more than two of these methods at a significance of $P < 0.05$. Any recombinant regions were removed for subsequent analysis.

5.3.3 Phylogenetic analysis

To reconstruct the spread of FIV_{Lru} lineages across the coastal southern Californian landscape, a Bayesian phylogeographic analysis of viral diffusion in continuous space was

conducted using Bayesian Evolutionary Analysis Sampling Trees (BEAST) version 1.10 (Drummond et al., 2012). Tree tip dates were specified according to sample date, with sampling location (latitude and longitude) specified as a continuous trait. Prior to incorporating phylogeographic random walk models, we tested several models comprising different substitution models and molecular clocks and selected the most appropriate using marginal likelihood estimation with path and stepping-stone sampling (Baele et al., 2012). The highest supported model included the HKY substitution model with gamma distributed rate heterogeneity and a proportion of invariant sites, the two-partition codon partition model, and a lognormal uncorrelated relaxed molecular clock. Using these parameters, relaxed random walk (RRW) models (Cauchy RRW, gamma RRW, and lognormal RRW) were then tested against a model assuming no dispersal rate variation among branches (Brownian random walk). The gamma RRW model was chosen, having significantly higher support than lognormal RRW and Brownian models, and marginally higher support than Cauchy RRW (see Appendix Table A5.1.1 for all model selection results). We ran three sets of 100 million Markov chain Monte Carlo iterations, sampling every 10,000 iterations and excluding the initial 10% of each set as burn-in. Model convergence and effective sample size (minimum 200) was checked and parameters evaluated using Tracer version 1.7 (Rambaut et al., 2018). A maximum clade credibility tree was constructed from the sampled trees using TreeAnnotator version 1.10 and visualised using FigTree version 1.4.3. SPREAD3 was used to visualise the spatiotemporal distribution of viral lineages based on the maximum clade credibility tree (Figure 5.1). Phylogenetic relatedness among each pair of FIV_{Lru} isolates was estimated by calculating patristic distances (the sum of branch lengths linking two nodes in a tree) based on the maximum clade credibility tree.

5.3.4 Landscape and host effects on FIV_{Lru} relatedness

To quantify the extent to which variation in patristic distances among FIV_{Lru} isolates is explained by host and landscape factors, we employed generalised dissimilarity modelling (GDM) using the *gdm* package in R (Fitzpatrick & Keller, 2015). GDM is a tool originally developed for community-level modelling of species turnover that performs matrix regressions to test associations between dissimilarities of predictor variables with dissimilarities of response variables among sample locations (Ferrier & Guisan, 2006). GDM allows non-linear responses and accounts for geographic distance among samples, and has recently been adapted for landscape genetic (Fitzpatrick & Keller, 2015) and phylogeographic analysis (Fountain-Jones et al., 2017c). Matrix permutation (100 permutations) with backwards elimination was used to calculate model and variable significance and to estimate variable importance. During backward elimination, as non-explanatory variables were removed, the percent deviance explained by each successive model did not change; deviance explained only dropped as explanatory variables were removed. Thus, the top model was identified as the model with the highest deviance explained and containing the fewest variables. Uncertainty was estimated for each variable response by subsampling site-pairs across 100 bootstrap iterations.

To assess host and landscape factors that influence study-wide variation in FIV_{Lru} relatedness, we first analysed all FIV_{Lru} samples together (n = 51). In addition, our goal was to analyse each population individually to account for among-host population differences in host or landscape factors affecting FIV_{Lru} relatedness. Sample size limitations necessitated that we combine some of the host populations into two broader groups – one comprising all FIV_{Lru} isolates collected northwest of Los Angeles (populations North-101, South-101, and East-405; n = 26), and another comprising all FIV_{Lru} isolates collected southeast of Los Angeles (populations East-5 and West-5; n = 25). We were unable to include host population

as a predictor because GDM measures dissimilarity among continuous or binary variables and is unable to analyse non-binary categorical variables. However, because host population designations were based on host genetic structure (Kozakiewicz et al. *in review*), we considered the inclusion of host relatedness in our analysis to largely account for any effect of host population on FIV_{Lru} patristic distance.

We quantified landscape factors both *at* and *between* sample locations to investigate how these factors explain FIV_{Lru} phylogenetic relatedness. When tested against patristic distances within the GDM framework, each of these two categories of landscape variable assessed different ways in which landscape heterogeneity can drive genetic processes. Factors quantified at sample locations ("site-based") investigated isolation by environment; that is, where genetic distance is determined by dissimilarity in local environmental factors due to processes such as natal habitat preference induction (in this case, of hosts) or local adaptation (Wang & Bradburd, 2014). Factors quantified between sample locations ("resistance-based") investigated isolation by resistance; that is, where genetic distance is determined by the extent to which the landscape intervening sample locations reduces functional connectivity (McRae, 2006).

Site-based landscape factors comprised a series of land cover types derived from the National Land Cover Database (NLCD; Homer et al., 2015) and quantified within buffers around sample locations. Buffers were sized according to published estimates of mean home range size for each sex within each of the three sampling areas (northwest LA: males = 7.1 km^2, females = 3.5 km^2; southeast LA: males = 6.8 km^2, females = 4.4 km^2; San Diego: males = 5.3 km^2, females = 3.8 km^2; Riley et al., 2010). The relative proportions of each urban (open, low, medium, and high density) and non-urban (forest, scrub, and grassland) land cover type were quantified within each buffer using Geospatial Modelling Environment (www.spatialecology.com). We also calculated the distance of each capture location to the

nearest urban edge using the Near tool in ArcGIS 10.3 (ESRI). Urban edge was defined from the NLCD impervious surface layer by calculating for each 30 m by 30 m raster cell whether a majority of cells within a 1 km radius had any value of impervious surface greater than zero, with cells meeting this criterion considered within the urban edge.

Resistance-based landscape factors were quantified using a landscape resistance framework. Landscape factors predicted have either positive or negative effects on bobcat (and thus FIV_{Lru}) connectivity were used to construct a series of resistance surfaces that each reflected these predictions. Factors predicted to positively influence connectivity were topographic roughness, vegetation density, and streams, while urbanisation (measured as percentage impervious surface) and roads were predicted to negatively influence connectivity. Further details of resistance surface parameterisation, ecological justifications of predictions, and landscape data sources are provided in Kozakiewicz *et al.* (*in review*). Circuitscape 4.0.5 (McRae et al., 2008) was used to model connectivity between sample locations along each resistance surface to produce pairwise matrices that measured the predicted effect of a given landscape factor on connectivity among each pair of sample locations. The effect of geographic distance was included in all models

Pairwise host relatedness was calculated as the inverse of the proportion of shared alleles across 13,520 single nucleotide polymorphism loci generated using double-digest restriction-site-associated DNA sequencing (see Kozakiewicz et al. *in review* for details). Host sex was coded as an integer (0 = female, 1 = male). All landscape and host variables were included as predictors of FIV_{Lru} patristic distance in GDM. To account for potential temporal autocorrelation (i.e., sequences sampled at the same time may be more similar), sample date as decimal years after the earliest sample date (30 September 1996) was included as an additional predictor. Prior to analysis, landscape resistance, host relatedness, and FIV_{Lru}

patristic distance matrices were converted to dissimilarity matrices, while all other variables were rescaled to values between zero and one.

5.3.5 Analysis of FIV_{Lru} spatial spread

We quantified rates of FIV_{Lru} spread across the landscape and tested the influence landscape factors. All analysis of FIV_{Lru} spread was conducted using *Seraphim* (Dellicour et al., 2016a) in R. One thousand trees were randomly sampled from our BEAST phylogeographic analysis and for all branches in each tree we extracted dates and locations associated with the start and end of each branch. Using this spatiotemporal information, we estimated dispersal statistics that describe the median velocity of FIV_{Lru} dispersal (in km/year) and the mean weighted diffusion coefficient (defined as diffusivity or the rate of spread in km^2/year; Trovão et al., 2015).

We tested the effects of five heterogeneous landscape factors – impervious surface, vegetation density, topographic roughness, streams, and roads on dispersal velocities associated with phylogenetic branches, using the resistance surfaces described above. *Seraphim* implements least-cost path analysis (Adriaensen et al., 2003) to estimate environmental distances among branch start and end locations for each resistance surface. Branch durations were linearly regressed against associated least-cost path distances for each landscape factor while accounting for the effect of isolation by distance alone (null model) to obtain a regression coefficient, R, and a modified coefficient of determination, Q. $R > 0$ indicates that the effect of environmental distance on branch duration is positive, and $Q > 0$ indicates that the effect of environmental distance on branch duration is greater than that of the null model. Values for R and Q were calculated for each of the 1000 sampled trees to obtain distributions of each. Landscape factors for which >90% of R and Q were distributed

above zero were considered potentially explanatory. Significance of *Q* distributions for potentially explanatory landscape factors was estimated using Bayes factor support calculated by randomisation of branch locations for each sampled tree and interpreted according to the guidelines of Kass and Rafferty (1995). Only branches with start dates after 1980 were included in tests of landscape effects to reduce temporal mismatch between estimated dispersal velocities and landscape factors (which were commensurate with sample collection dates).

As with the analysis of FIV_{Lru} relatedness above, we repeated all analyses of FIV_{Lru} spatial spread for the northwest of Los Angeles and southeast of Los Angeles regions separately to investigate differences among regions in rates and landscape drivers of FIV_{Lru} spread. Subsets of extracted branches were compiled for each region according to start and end locations of each branch and analysed separately as above.

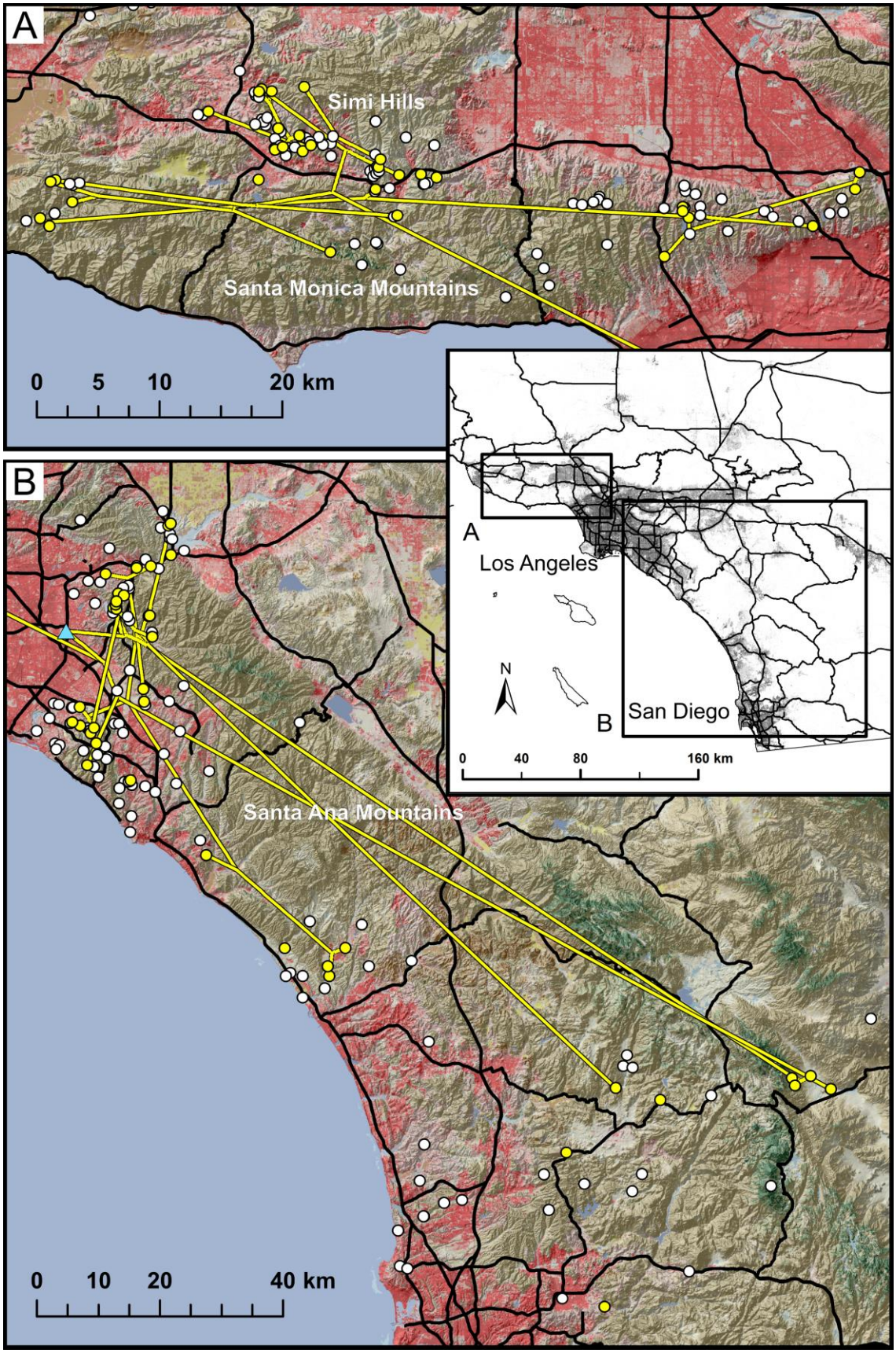

Figure 5.1. Spatiotemporal dispersal of feline immunodeficiency virus (FIV_{Lru}) in bobcats, reconstructed using continuous Bayesian phylogeographic analysis. Yellow lines indicate bifurcating branches showing phylogenetic relationships, originating from an ancestral node

shown as a light blue triangle. Circles indicate bobcat sample locations, with FIV_{Lru}-positive samples shown in yellow, and FIV_{Lru}-negative samples shown in white. Areas northwest (A) and southeast (B) of Los Angeles are shown, with a single branch joining the two areas. See Appendix Figure A5.1.1 for additional panels showing close-up views of phylogeographic relationships.

5.4 Results

5.4.1 Analysis of factors affecting FIV_{Lru} relatedness

The GDM analysis for host and landscape factors affecting FIV_{Lru} phylogenetic relatedness across the entire coastal southern California study region indicated greatest support for the model containing site-based landscape factors comprising geographic distance, scrub land cover, impervious surface resistance distances, forest land cover, host relatedness, and vegetation density resistance distances (Figure 5.2a). Geographic distance explained the greatest percentage of model deviance (8.31%), with infected bobcats sampled further apart from one another having more distantly related FIV_{Lru} ($P < 0.01$). Scrub land cover explained 3.01% of model deviance, with infected bobcats sampled at locations that differ more in the amount of scrub land cover having more distantly related FIV_{Lru} ($P = 0.05$). Impervious surface resistance distances explained 2.45% of model deviance, with greater amounts of impervious surface between bobcat sample locations having more distantly related FIV_{Lru} ($P = 0.03$). Overall, this model explained 52.5% of variation in FIV_{Lru} patristic distance.

Within the region northwest of Los Angeles, the best supported model contained forest land cover, geographic distance, grassland land cover, and host relatedness (Figure 5.2b). Forest land cover explained the greatest percentage of model deviance (31.9%), with infected bobcats sampled at locations that differ more in the amount of forest land cover having more distantly related FIV_{Lru} ($P = 0.05$). Geographic distance explained 7.26% of

model deviance, with infected bobcats sampled further apart from one another having more distantly related FIV_{Lru} ($P < 0.01$). Overall, this model explained 33.1% of variation in FIV_{Lru} patristic distances.

Southeast of Los Angeles, most of the tested models were non-significant, but the best supported model ($P = 0.04$) contained scrub land cover, geographic distance, vegetation density resistance distances, and topographic roughness resistance distances (Figure 5.2c). Scrub land cover explained the greatest percentage of model deviance (31.6%), with infected bobcats sampled at locations that differ more in the amount of scrub land cover having more distantly related FIV_{Lru} ($P = 0.05$). Geographic distance explained 27.2% of model deviance, with infected bobcats sampled further apart from one another having more distantly related FIV_{Lru}; however, this effect was only marginally significant ($P = 0.07$). Overall, this model explained 17.7% of variation in FIV_{Lru} patristic distance.

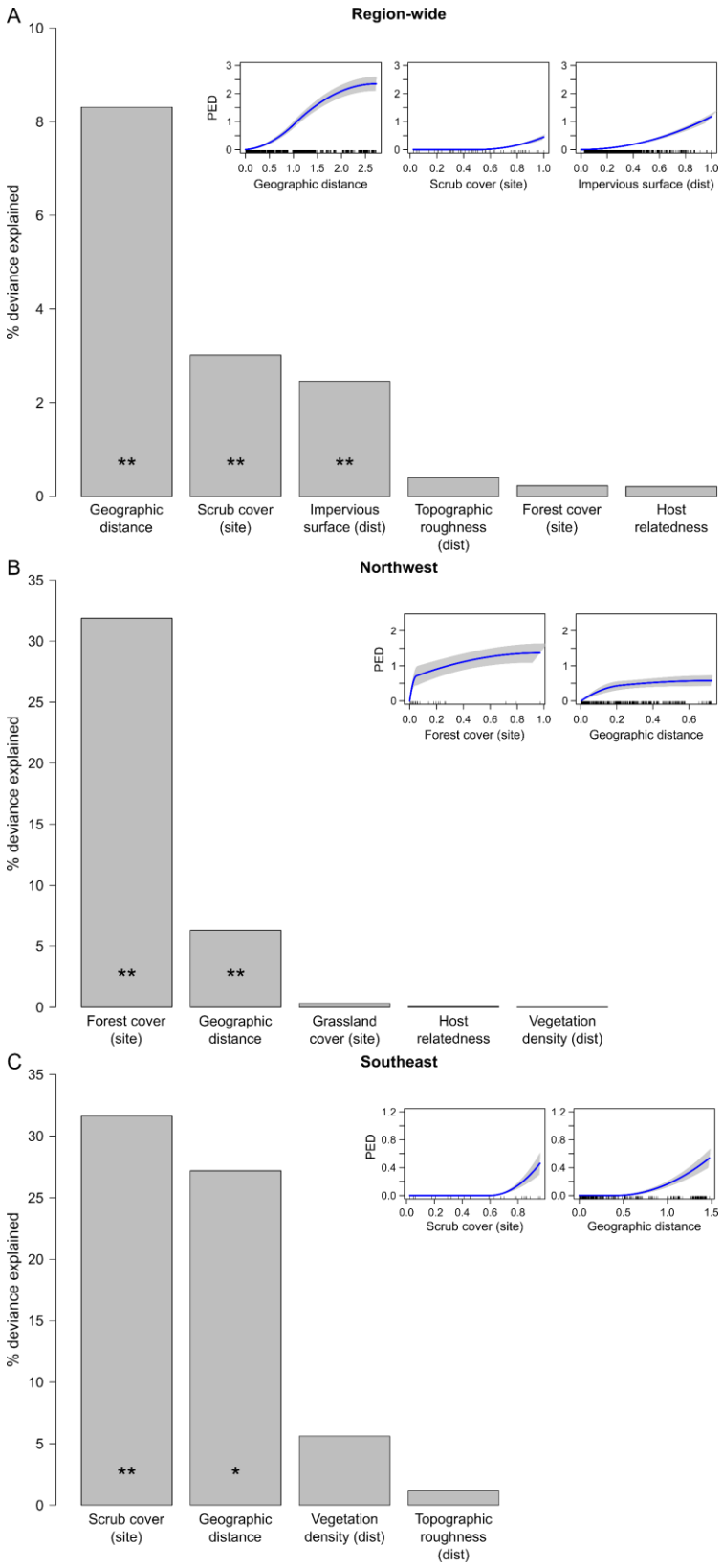

*Figure 5.2. Generalised dissimilarity models showing factors affecting FIV_{Lru} phylogenetic differentiation in bobcats across the coastal southern California region (A), and in more localised areas northwest (B) and southeast (C) of Los Angeles. **Significant effects (P < 0.05), *marginally significant effects (P < 0.10). Partial ecological distance (PED) plots*

show the fitted I-splines for each significant environmental variable, indicating the rate of change in FIV_{Lru} patristic distance as environmental dissimilarity increases.

5.4.2 Analysis of factors affecting rates of FIV_{Lru} spread

Dispersal velocities are the linear rate at which individual phylogenetic branches move across the landscape. The median weighted dispersal velocity of FIV_{Lru} branches across the entire coastal southern California region was 0.806 km/y (95% HPD: 0.467–1.38). The median dispersal velocity of FIV_{Lru} was marginally higher northwest of Los Angeles (0.856 km/y; 95% HPD: 0.489–1.40) than in the southeast (0.735 km/y; 95% HPD: 0.434–1.27). The weighted diffusion coefficient is a cumulative estimate of branch dispersal velocities that measures the rate of expansion of the area into which viral lineages have spread or "diffused". The weighted diffusion coefficient for the entire southern California region was 11.75 km^2/year (95% HPD: 5.99–21.1). The diffusion coefficient was substantially higher southeast of Los Angeles (12.5 km^2/y; 95% HPD: 6.08–23.76) than in the northwest (2.80 km^2/y; 95% HPD: 1.50–5.00). These results suggest that while individual FIV_{Lru} dispersal velocities are minimally different northwest and southeast of Los Angeles, the overall rate at which lineages spread throughout the landscape was greater southeast of Los Angeles.

Northwest of Los Angeles, no factors explained FIV_{Lru} phylogeographic branch dispersal velocities better than the null model, with the single factor having the second greatest explanatory power being vegetation density (proportion of $Q > 0 = 0.61$; BF = 2.98). Southeast of Los Angeles, riparian habitat had greater explanatory power than the null model, but this effect had little Bayes factor support (proportion of $Q > 0 = 0.95$, BF = 1.63). Thus, we found little effect of landscape on dispersal velocities in each of the northwest and southeast areas. Region-wide, vegetation density had greater explanatory power than the null model and had positive Bayes factor support (proportion of $Q > 0 = 0.70$, BF = 3.42). This

relatively low proportion of $Q > 0$ suggests a relatively weak positive effect of vegetation density on dispersal velocity that is not present in all the sampled phylogeographic trees but is significant based on Bayes factor estimation.

5.5 Discussion

Our wide distribution of samples enabled us to quantify how host and landscape factors influencing FIV_{Lru} dynamics vary among local areas and among local (i.e., either northwest or southeast of Los Angeles) and region-wide (i.e., study-wide) spatial scales. Patterns of FIV_{Lru} phylogenetic relatedness at the regional scale, as well as locally within both the northwest of Los Angeles and southeast of Los Angeles groups, were driven by isolation by distance patterns, evidenced by the significant effects of geographic distance on FIV_{Lru} phylogenetic relatedness. However, landscape factors explained some variation in FIV_{Lru} phylogenetic relatedness. At the regional scale, we found that bobcats sampled from sites that were more divergent in the amount of scrub cover had more distantly related FIV_{Lru} strains. We also found that resistance distances representing a negative effect of impervious surface on connectivity were associated with greater differentiation among FIV_{Lru} strains regionally. At a local scale southeast of Los Angeles, the positive effect of divergence in the amount of scrub cover on FIV_{Lru} phylogenetic differentiation persisted; however, northwest of Los Angeles, divergence among sites in forest landcover resulted in increased FIV_{Lru} phylogenetic differentiation. Although we found no effect of landscape factors on FIV_{Lru} dispersal velocities in areas northwest or southeast of Los Angeles individuals, we found a weak positive effect of vegetation density on dispersal velocities at the regional scale.

In our study, urban development (measured as percentage impervious surface) had a negative effect on FIV_{Lru} phylogenetic relatedness at a region-wide scale. Urbanisation is an

important factor limiting bobcat connectivity in coastal southern California (Kozakiewicz et al. *in review*; Riley et al., 2006; Ruell et al., 2012), and has been shown to reduce FIV_{Lru} gene flow at smaller spatial scales (Fountain-Jones et al., 2017a). This region-wide effect of urbanisation likely reflects the major impediment that the highly urbanised Los Angeles Basin represents for wildlife connectivity. However, geographic distance, and either forest or scrub land cover, were the most important predictors both regionally and within the areas northwest and southeast of Los Angeles. This may reflect a reduced effect of urbanisation on FIV_{Lru} connectivity relative to that of its host, perhaps due to the need for a dispersing bobcat to establish a territory and successfully reproduce for gene flow to occur, whereas an FIV_{Lru} strain could in theory more readily disperse across the landscape through host movement, contact, and transmission. Such a divergence in factors influencing host and pathogen gene flow may also explain why we did not detect any effect of host relatedness.

We found evidence that bobcats sampled from habitats that differed more in the amount of natural habitat had more distantly related FIV_{Lru} relative to bobcats sampled from areas with similar amounts of natural habitat. This effect was particularly pronounced northwest of Los Angeles, where even small differences in forest landcover play a major role in explaining FIV_{Lru} phylogenetic differentiation. Southeast of Los Angeles, as well as at the region-wide scale, we found a similar effect of scrub whereby samples from sites that were more divergent in the amount of scrub cover tended to have more distantly related FIV_{Lru} strains. Such patterns indicate an isolation by environment (IBE) effect, whereby phylogenetic differentiation and environmental differentiation are positively correlated (Wang & Bradburd, 2014). IBE can occur through a number of mechanisms that result in decreased gene flow among different environments, including natural selection against immigrants, sexual selection against immigrants, and habitat-biased dispersal (Wang & Bradburd, 2014). Forest habitat in coastal southern California is sparse (maximum proportion

of forest landcover at any sample location was 0.2), meaning that even bobcats in areas that have relatively high amounts of forest habitat must frequently make use of other habitat types, usually scrub. This necessity for bobcats to utilise habitat types other than forest makes it unlikely that variation in forest habitat is imposing selective pressure on bobcats or FIV_{Lru}.

We believe the most likely explanation for the observed IBE patterns is reduced exchange of FIV_{Lru} strains among areas that differ in the amount of natural habitat present, which may reflect reduced movement of bobcats among urban and non-urban areas. Scrub land cover southeast of Los Angeles is negatively correlated with total combined urban landcover (R = -0.71), and forest land cover northwest of Los Angeles is positively correlated with distance to urban edge (R = 0.77; see Appendix Tables A5.1.2–A5.1.4 for all land cover correlations). Thus, although we did not find a direct effect of urban landcover on FIV_{Lru} phylogenetic differentiation northwest or southeast of Los Angeles, the effects of urbanisation on the distribution of natural habitat may explain the patterns we observed. This explanation is supported by Fountain-Jones *et al*. (2017a), who found that FIV_{Lru} gene flow is greater with increasing distance from urban development. This would suggest a decreased likelihood that individuals would disperse from a natural area to a more urbanised area, instead of to a similarly natural area. Although some individual examples of dispersal (with successful reproduction) among urban and natural areas have been documented (Riley et al., 2010), pile-up of bobcat home ranges near roads and the urban edge may reduce the probability of dispersing individuals establishing territories in urban areas (Riley et al., 2006). This may further reduce the frequency of pathogen exchange among urban and non-urban areas.

The relationship between forest landcover and FIV_{Lru} phylogenetic differentiation that we found northwest of Los Angeles must also be considered in the context of population structure in this area. The northwest area comprises three genetically distinct bobcat

populations, separated by major highways (Kozakiewicz et al. *in review*; Riley et al., 2006; Serieys et al., 2015), which in turn drive FIV_{Lru} phylogeographic structure (Kozakiewicz et al. *in prep*). The already patchily distributed forest landcover is largely limited to the populations in the Santa Monica Mountains, whereas the area to the north in the Simi Hills is virtually devoid of forest vegetation (natural vegetation here is almost entirely grassland and scrub). It is possible that this population genetic differentiation, coupled with the stark difference among populations in the amount of forest landcover, is at least in part driving the observed relationship. Ideally, each of these FIV_{Lru} populations would be analysed separately, as was conducted for the host populations by Kozakiewicz *et al*. (*in review*), to account for the effect of population structure. However, FIV_{Lru} sample size constraints did not permit such a design in this instance.

We also recognise that deeper phylogenetic relationships among FIV_{Lru} lineages may be a stronger predictor of FIV_{Lru} phylogenetic differentiation than landscape heterogeneity. In contrast to their hosts, FIV_{Lru} strains do not exchange genes with every generation (although recombination does occur every so often; Lee et al., 2014), and thus multiple distantly related FIV_{Lru} lineages can circulate independently within the same host population. This can result in very high phylogenetic differentiation among strains within a small spatial area that reflect deeper phylogenetic relationships more so than any contemporary factors influencing connectivity. Thus, the large phylogenetic differences among lineages are, for example, equivalent to genetic differences among multiple sympatric, reproductively isolated species and do not directly reflect contemporary gene flow patterns. Particularly southeast of Los Angeles, phylogenetic relationships among spatially proximate lineages were relatively deep (Kozakiewicz et al. *in prep*), and may confound our ability to detect effect of landscape on FIV_{Lru} gene flow. Thus, our models for predicting FIV_{Lru} phylogenetic differentiation had approximately half the explanatory power southeast of Los Angeles compared to the

northwest (which had shallower phylogenetic relationships). However, we did detect site-based landscape effects on FIV_{Lru} phylogenetic differentiation, indicating a tendency of one lineage or another to occupy certain environmental niches.

Overall, rates of FIV_{Lru} dispersal were relatively slow (0.43–1.4 km/year) compared with rates for other host-pathogen systems (Dellicour et al., 2017). This effect may be due to urbanisation, whereby higher rates of FIV_{Lru} dispersal might be expected in less urban areas where bobcat home ranges are larger (Riley et al., 2010). For example, rabies dispersal in urban dogs was 0.65 km/year compared with up to 22 km/year in non-urban areas (Bourhy et al., 2016; Dellicour et al., 2017). Further supporting the importance of natural habitat in facilitating FIV_{Lru} transmission, we found a weak positive effect of vegetation density on dispersal velocities at the region-wide scale. However, none of the tested landscape factors were associated with FIV_{Lru} dispersal velocity either northwest or southeast of Los Angeles.

While we observed no appreciable difference in FIV_{Lru} dispersal velocities among our northwest and southeast of Los Angeles study areas, FIV_{Lru} diffusion coefficients northwest of Los Angeles were approximately half the rate as that southeast of Los Angeles, suggesting FIV_{Lru} spread is more constrained in this area. The host populations northwest of Los Angeles are relatively isolated compared to those in the southeast (Kozakiewicz et al. *in review*), resulting in decreased opportunity for long-distance movements. In contrast, southeast of Los Angeles, we observed three FIV_{Lru} lineages that dispersed distances upwards of 100 km, facilitated by the lack of any substantial barrier to bobcat connectivity to the east (Kozakiewicz et al. *in review*).

Overall, our results suggest that urbanisation plays less of a role in constraining FIV_{Lru} transmission than it does to limit bobcat connectivity. However, we found evidence that the degree of relatedness among pathogen isolates was reduced among areas that were more

divergent in the amount of natural habitat available, potentially also related to urbanisation. We also demonstrate differences among host populations in the natural landscape factors that influence these patterns of pathogen relatedness. Our results suggest that in the event of a severe disease outbreak, different populations may need to be managed not only as distinct units, but as units among which the factors driving disease transmission may differ. Further, even within a single host population, the management of disease transmission in more urbanised areas may need to be considered separately from transmission in natural areas. This work demonstrates the utility of emerging ecological phylogenetic methods in elucidating the factors shaping pathogen transmission in urban landscapes. Due to the broad distribution of bobcats and FIV_{Lru} in North America across a variety of environments (Reding et al., 2012; Lagana et al., 2013; Carver et al., 2016), this system represents an ideal opportunity to investigate factors driving pathogen transmission in a variety of urban and non-urban contexts. With the continuing expansion of urban landscapes globally, understanding how urbanisation impacts patterns of disease will be increasingly valuable.

5.6 Appendix

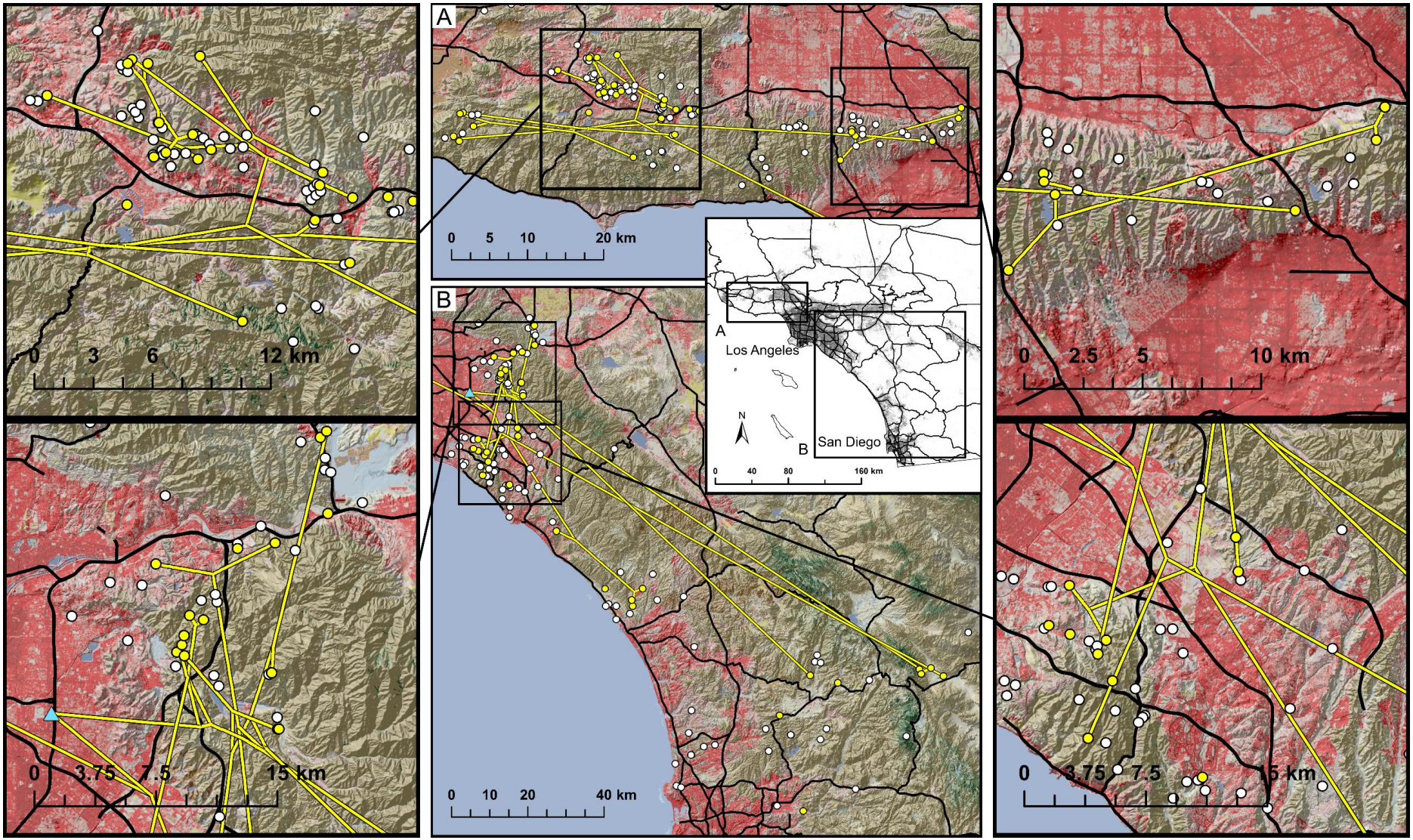
A
0 5 10 20 km
B
0 10 20 40 km
A
Los Angeles
B
San Diego
N
0 40 80 160 km
0 3 6 12 km
0 3.75 7.5 15 km
0 2.5 5 10 km
0 3.75 7.5 15 km

Figure A5.1.1. Spatiotemporal dispersal of feline immunodeficiency virus (FIV_{Lru}) in bobcats, reconstructed using continuous Bayesian phylogeographic analysis, as shown in Figure 5.1. Additional panels show phylogenetic relationships among spatially proximate FIV_{Lru} isolates.

Table A5.1.1. BEAST model selection results for tests of substitution models, molecular clock, and random walk models. Log-maximum likelihood scores for path sampling and stepping-stone sampling are shown.

Substitution model	Molecular clock	Random walk model	Log ML (PS)	Log ML (SS)
GTR	strict	N/A	-8457.96	-8457.72
GTR	uncorrelated relaxed	N/A	-8354.26	-8354.1
HKY	strict	N/A	-8477.64	-8477.39
HKY	uncorrelated relaxed	N/A	-8351.76	-8352.04
HKY	uncorrelated relaxed	Brownian	-8192.76	-8192.97
HKY	uncorrelated relaxed	Lognormal	-8185.53	-8185.9
HKY	uncorrelated relaxed	Cauchy	-8182.57	-8182.57
HKY	uncorrelated relaxed	Gamma	-8181.8	-8181.98

Table A5.1.2. Pearson correlation coefficients (R) among land cover variables measured at bobcat sample locations region-wide

	Developed, open	Developed, low	Developed, medium	Developed, high	Developed, total	Dist. to urban edge	Forest	Scrub
Developed, low	0.67							
Developed, medium	0.57	0.88						
Developed, high	0.36	0.33	0.54					
Developed, total	0.80	0.89	0.92	0.66				
Dist. to urban edge	-0.58	-0.45	-0.38	-0.20	-0.50			
Forest	-0.33	-0.32	-0.26	-0.10	-0.31	0.44		
Scrub	-0.66	-0.72	-0.73	-0.40	-0.77	0.24	0.39	
Grassland	-0.09	-0.14	-0.10	-0.11	-0.13	0.26	-0.37	-0.45

Table A5.1.3. Pearson correlation coefficients (R) among land cover variables measured at bobcat sample locations northwest of Los Angeles

	Developed, open	Developed, low	Developed, medium	Developed, high	Developed, total	Dist. to urban edge	Forest	Scrub
Developed, low	0.52							
Developed, medium	0.38	0.85						
Developed, high	0.39	0.21	0.55					
Developed, total	0.72	0.83	0.90	0.69				
Dist. to urban edge	-0.75	-0.67	-0.57	-0.26	-0.71			
Forest	-0.38	-0.45	-0.35	-0.08	-0.40	0.77		
Scrub	-0.68	-0.80	-0.78	-0.43	-0.86	0.75	0.48	
Grassland	-0.06	-0.04	-0.04	-0.15	-0.09	-0.24	-0.47	-0.37

Table A5.1.4. Pearson correlation coefficients (R) among land cover variables measured at bobcat sample locations southeast of Los Angeles

	Developed, open	Developed, low	Developed, medium	Developed, high	Developed, total	Dist. to urban edge	Forest	Scrub
Developed, low	0.80							
Developed, medium	0.73	0.93						
Developed, high	0.36	0.51	0.54					
Developed, total	0.86	0.95	0.94	0.66				
Dist. to urban edge	-0.59	-0.45	-0.38	-0.26	-0.50			
Forest	-0.36	-0.23	-0.19	-0.16	-0.28	0.57		
Scrub	-0.66	-0.67	-0.69	-0.39	-0.71	0.15	0.33	
Grassland	-0.07	-0.18	-0.13	-0.07	-0.13	0.34	-0.28	-0.53

Chapter 6

General discussion

Urbanisation is a key factor driving habitat fragmentation and loss, reducing connectivity in wildlife populations and driving changes in pathogen prevalence and dynamics (Bierwagen, 2007; Bradley & Altizer, 2007; Brearley et al., 2013). Urban areas are expanding globally, with the human urban population projected to increase by 60% on current figures to 6.7 billion by 2050 (United Nations, 2018). Therefore, research examining the impacts of urban development on wildlife population connectivity and pathogen dynamics is crucial to predicting and managing these impacts into the future. By implementing a range of contemporary tools for integrating molecular data from hosts and pathogens with measurements of landscape heterogeneity, the work presented in this thesis contributes to our knowledge of factors driving host connectivity and pathogen dynamics in an urban system.

In this thesis, I focused on populations of bobcat (*Lynx rufus*) in coastal southern California and a directly transmitted pathogen, feline immunodeficiency virus (FIV_{Lru}). Following an initial literature review exploring the use of landscape genetics in disease ecology (Chapter 2; Kozakiewicz et al., 2018a), I conducted a landscape genomic study of bobcats, identifying five populations isolated by major highways and a negative effect of urban landcover on bobcat connectivity both within and among populations (Chapter 3; Kozakiewicz et al. *in review*). I then performed a phylogeographic analysis of FIV_{Lru}, finding that FIV_{Lru} phylogeographic structure generally reflected patterns of host population structure, but that there was substantial variation among populations in the genetic diversity of FIV_{Lru} strains present, reflecting differences in host effective population size (Chapter 4; Kozakiewicz et al. *in prep*). Finally, I implemented ecological phylogenetic tools to identify

host and fine-scale landscape factors affecting FIV_{Lru} phylogeography and velocities of FIV_{Lru} lineage spread across the landscape. I identified natural vegetation as an important factor influencing pathogen spread, with potentially reduced rates of transmission among urban and natural areas (Chapter 5; Kozakiewicz et al. *in prep*). In summary, this work:

i) highlights a variety of emerging methods for elucidating the effects of landscape heterogeneity on host and pathogen spatial genetic structure

ii) identifies factors affecting bobcat connectivity in a highly urbanised environment

iii) indicates how host population structure and landscape heterogeneity shape FIV_{Lru} phylogenetic structure and transmission dynamics

iv) demonstrates the utility of incorporating multiple populations and spatial scales into analyses of genetic structure to contextualise observed patterns and relationships

6.1 Findings

The high resolution of contemporary molecular analyses provides a means of linking gene flow in wildlife and their pathogens to fine scale heterogeneous landscape factors, a task that has long been a challenge for ecologists. For free-living organisms, the field of landscape genetics is now widely implemented and has greatly increased our understanding of how wildlife connectivity is shaped by the landscape (Manel et al., 2003; Storfer et al., 2010; Manel & Holderegger, 2013). However, in Chapter 2, I identified that the uptake of landscape genetics and similar methods for the study of pathogens has been slow. In my review of this literature, I emphasised the utility of landscape genetics for disease ecology and detailed how existing landscape genetic tools may be used to elucidate pathogen dynamics. Further, I

highlighted emerging frontiers, including multi-species landscape genetics frameworks (Leo et al., 2016; Schwabl et al., 2017) and ecological phylogenetic approaches (Dellicour et al., 2016b; Fountain-Jones et al., 2017c) that I anticipate will advance our understanding of factors influencing pathogen dynamics in heterogeneous landscapes.

Previous studies have provided valuable foundational knowledge regarding habitat use and the influence of major roads on connectivity among bobcat populations in coastal southern California (Riley et al., 2003, 2006; Ordeñana et al., 2010; Lee et al., 2012; Ruell et al., 2012; Poessel et al., 2014; Serieys et al., 2015; Fountain-Jones et al., 2017a). However, the factors influencing connectivity within populations were still largely unknown. By analysing samples from across the entire region under a unified landscape genomic framework, I achieved an in-depth yet broad investigation of factors influencing bobcat gene flow both within and among populations, which varied in spatial area and landscape composition. I revealed that landscape effects on gene flow were more common at a region-wide spatial scale, with urban landcover and topographic roughness having negative and positive effect on gene flow, respectively. The negative effect of urban landcover on gene flow persisted within populations irrespective of size and overall amounts of urban development, suggesting a pervasive effect of urbanisation largely independent of spatial scale. However, I found that the strongest negative effect of urbanisation on gene flow was in a population where riparian habitat had largely been lost to development.

Many pathogens rely on their hosts for movement across the landscape. However, host and parasite genetic structure are frequently divergent (Mazé-Guilmo et al., 2016). For example, there is evidence of FIV_{Lru} transmission across a major freeway southeast of Los Angeles despite this road being a major impediment to host gene flow (Lee et al., 2012; Fountain-Jones et al., 2017a). I reconstructed phylogenetic relationships among FIV_{Lru} isolates sampled from across coastal southern California to assess the permeability of several

major barriers to bobcat dispersal to FIV_{Lru} spread. This analysis provided temporal estimates of FIV population divergence over the past 118 years, which coincides with a period of rapid urban growth, and demonstrates losses in connectivity over this time. I found that FIV_{Lru} populations northwest of Los Angeles exhibited strong phylogeographic structure in accordance with host population, whereas populations southeast of Los Angeles had relatively weak phylogeographic structure. However, rates of FIV_{Lru} transfer across barriers did not differ among these areas and were low overall. High effective population sizes, which I observed for host and pathogen southeast of Los Angeles, can contribute to incomplete lineage sorting (Pamilo & Nei, 1988; Maddison, 1997; Pease & Hahn, 2013) that can cause incongruence among host and parasite genetic structure (e.g., Symula et al., 2011; Torres-Pérez et al., 2011). Thus, incomplete lineage sorting, and not frequent road-crossing, may explain the lack of FIV_{Lru} phylogeographic structure southeast of Los Angeles relative to host population structure.

Urban development has been shown to affect patterns of disease prevalence (Brearley et al., 2013). However, the mechanisms driving these changes are often poorly understood owing to the difficulty of relating transmission events to specific landscape factors (Craft & Caillaud, 2011; Gilbertson et al., 2018). By testing the influence of specific heterogeneous landscape and host factors on patterns of FIV_{Lru} phylogenetic relatedness and velocities of FIV_{Lru} lineage spread, I found that natural habitat appeared to be an important factor shaping FIV_{Lru} dynamics. Specifically, increasing differentiation among FIV_{Lru} isolates was observed as differences among sample locations in the amount of natural habitat available increased. This suggests that FIV_{Lru} transmission predominates within more urban areas and within more natural areas, with more limited FIV_{Lru} transmission between urban and natural areas. Interestingly, forest habitat was most important for FIV_{Lru} transmission northwest of Los Angeles, while scrub habitat was most important in the southeast. Phylogenetic

differentiation among FIV_{Lru} populations northwest and southeast of Los Angeles was explained by the heavily urbanised Los Angeles Basin. When evaluating the spatial dispersal velocity of FIV_{Lru}, I found a weak positive association between vegetation density and FIV_{Lru} velocity at a broader region-wide spatial scale. However, no fine scale (i.e. local to either the northwest or southeast areas) effect of landscape on dispersal velocity was observed.

6.2 Management implications

Coastal southern California is one of the most urbanised regions in North America, yet it is a global hotspot for biodiversity and endemism, with high concentrations of threatened species in remaining natural areas (Dobson et al., 1997; Myers et al., 2000). Increasing urbanisation here (and globally) will likely result in further losses of wildlife connectivity and changes in patterns of disease, both of which will have implications for species persistence (McCallum & Dobson, 2002; Smith et al., 2006; Bierwagen, 2007; Faust et al., 2018). Because bobcats are known indicators of connectivity in this region (Crooks, 2002), my work provides valuable insights into how patterns of urbanisation can impact the movement of wildlife and disease. I demonstrated that urban landcover can reduce bobcat connectivity even in areas where relatively little urbanisation is present. One population where riparian vegetation had largely been lost to development was suffering a severe loss of connectivity due to urbanisation, which places this population at heightened risk of extirpation. Riparian zones have shown to be heavily utilised by wildlife and may act as dispersal corridors, particularly where natural habitat is scarce (Machtans et al., 1996; Hilty & Merenlender, 2004; Gillies & St. Clair, 2008; LaRue & Nielsen, 2008). Thus, the preservation of riparian corridors may mitigate the negative impacts of urbanisation on connectivity in coastal southern California and should be prioritised.

My results also show that factors influencing wildlife connectivity and disease transmission may differ among populations, and that different management strategies may be required for different populations. In addition, although FIV_{Lru} is present regardless of levels of urban isolation, my results suggest reduced pathogen transmission between urban and natural areas. This could be due to increased FIV_{Lru} transmission further from urban areas, as has been suggested by Fountain-Jones *et al.* (2017a), which may reduce the relative likelihood of dispersal from natural areas to more urban areas. An alternative or concurrent mechanism might be home-range pile up near the urban edge, which may reduce the probability of dispersing individuals establishing territories in urban areas (Riley et al., 2006). Given limited pathogen exchange among urban and natural areas, separate management strategies for each of these environments may be beneficial.

The ability to predict patterns of disease across the landscape can provide invaluable guidance for interventions such as vaccination or culls targeting key areas during disease outbreaks, or to inform broader landscape management decisions for mitigating the risk of outbreaks in the first place. Oftentimes, measurements of host connectivity are used to inform these predictions (Kozakiewicz et al., 2018a). I found that despite a broad congruence between host and pathogen population structure, factors influencing pathogen transmission and spread at finer scales may differ from those influencing the host. For species-specific directly transmitted pathogens, such as FIV_{Lru}, it is often movement of non-reproducing hosts that can generate such incongruence, but for many other pathogens, factors such as additional host or vector species, environmental persistence, or pathogen reproductive mode can result in even greater divergence from host genetic structure (Mazé-Guilmo et al., 2016; Tesson et al., 2016). Thus, the inference of pathogen transmission using host genetic structure is likely to be unreliable in many instances. However, knowledge of host connectivity could help direct subsequent studies investigating pathogen transmission specifically. For example,

insights into bobcat connectivity provided in Chapter 3 provided valuable direction for my studies of FIV_{Lru} in Chapters 4 and 5. This thesis highlights a number of emerging landscape genetic and ecological phylogenetic approaches for linking pathogen transmission with landscape heterogeneity that should enhance the ability of managers to understand and predict patterns of disease spread in wildlife populations.

6.3 Future directions

Bobcats are among the most common and widely distributed mammals in North America, occupying a variety of habitats across a broad spectrum of climates and urban influence (Reding et al., 2012). The populations studied herein are among the most urbanised from across this species' range and have provided valuable insights into how urbanisation impacts host connectivity and pathogen transmission. However, it is unclear how applicable my findings may be in other parts of the species' range, and establishing "baseline" patterns of connectivity in more natural landscapes would help contextualise the patterns observed here in urban populations. The broad distribution of this species presents an excellent opportunity to assess factors affecting host connectivity and pathogen dynamics across a wide gradient of landscapes.

Approaches for linking pathogen dynamics and landscape heterogeneity to more complex systems, including multi-host and vector-borne diseases, remains a challenge. Even now, the work in this thesis is one of only a few examples integrating both host and pathogen genetic data in a landscape genetic framework (Kozakiewicz et al., 2018a). However, emerging approaches for integrating data from multiple hosts, vectors, and pathogens are promising to drive advancements in this area. Such approaches include explicit landscape

genetic frameworks for integrating multiple taxa (Leo et al., 2016; Schwabl et al., 2017), or community ecological methods adapted for genetic datasets (Fountain-Jones et al., 2017c).

The importance in landscape genetics of replication across multiple study areas and considering the effect of spatial scale on landscape genetic inference has been advocated for future studies (Cushman & Landguth, 2010a; Short Bull et al., 2011), but such approaches are rarely implemented. A recurrent theme of this thesis was to investigate landscape effects on genetic variation within multiple populations separately as well as at a broader scale that accounted for differentiation among populations. This replicated, multi-scale approach enabled robust estimation of factors influencing host connectivity and pathogen transmission and allowed me to tease apart local, population-specific influences from broader, more general patterns, and identify underlying causes of population-specific relationships. For example, I found a general pattern of urban landcover negatively influencing bobcat connectivity, but that this effect may be mitigated by the retention of riparian habitat. For FIV_{Lru}, I also found that in general, the amount of natural habitat available influenced phylogenetic relatedness, but that the type of habitat that was important varied among populations. These detailed insights would not have been possible if not for this replicate design, and my work strongly emphasises the value of this underutilised approach.

Next-generation sequencing (NGS) approaches for quantifying genome-wide variation are facilitating an increasing focus of landscape genomics on patterns of adaptation and selection (Storfer et al., 2018). Although I identified a small number of loci in bobcats that were potentially under selection, I did not explore this further due to the focus of my work on connectivity and gene flow. Because of the pressures imposed by urban development on wildlife, selection may be driving adaptive changes in response to the urban environment. For example, urban-associated bobcats tend to exhibit changes in home range size, habitat use, and prey species (Riley et al., 2003, 2006, 2010; Poessel et al., 2014). Such changes may

be driven in part by mortality risks associated with urban areas, such as vehicle collision, rodenticide exposure, and notoedric mange (Riley et al., 2007; Lyren et al., 2008b; Serieys et al., 2015; Fraser et al., 2018). Future projects could test for associations among individual loci and environmental variables and investigate links to specific genes affecting fitness to generate insights into the processes driving selection and adaptation in urban populations.

Advantages of NGS approaches could also be extended to the analysis of viruses. NGS approaches are less sensitive to factors such as mutations at primer sites, low virus copy number or sample degradation that prevented me from obtaining sequences for 20 FIV_{Lru}-positive samples out of a possible 73. In addition, each host carries a diversity of infectious agents, including different strains (or variants) of the same viruses. My approach sequenced the dominant FIV_{Lru} strain in each host, whereas NGS has the capacity to genotype entire populations of viruses within a single host and may also be used to concurrently genotype other potentially informative pathogens, including whole genomes of more slowly mutating DNA viruses and bacteria (Biek et al., 2015). Thus, NGS approaches would be better able to capture the overall diversity of pathogen populations and may increase our power to detect landscape and other factors influencing pathogen genetic variation.

6.4 Concluding remarks

In this thesis I have integrated landscape, host, and pathogen data to investigate how landscape heterogeneity influences host connectivity, and the subsequent effects on pathogen dynamics, in an urban environment. This thesis contributes substantially to our understanding of how increasing urbanisation impacts wildlife populations and patterns of disease transmission and spread. Few studies to date integrate genetic estimates of both host and pathogen connectivity in a heterogeneous landscape. My findings show that while host and

pathogen genetic structure is broadly congruent, the specific landscape factors driving fine-scale patterns of connectivity may differ among pathogens and their hosts. In addition, the factors influencing connectivity may differ among populations for both hosts and pathogens. Ongoing urban expansion globally will likely drive further reductions in host connectivity and alter patterns of disease transmission and spread. Continued development and implementation of methods linking host and pathogen genetic data with heterogeneous landscape factors is necessary to predict and manage the impacts of urbanisation on wildlife and their diseases.

References

Addis, BR, Lowe, WH, Hossack, BR, & Allendorf, FW. (2015). Population genetic structure and disease in montane boreal toads: more heterozygous individuals are more likely to be infected with amphibian chytrid. *Conservation Genetics*, ***16***, 833–844.

Adriaensen, F, Chardon, JP, De Blust, G, Swinnen, E, … Matthysen, E. (2003). The application of 'least-cost' modelling as a functional landscape model. *Landscape and Urban Planning*, ***64***, 233–247.

Allendorf, FW, Hohenlohe, PA, & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, ***11***, 697–709.

Anderson, CD, Epperson, BK, Fortin, MJ, Holderegger, R, … Spear, S. (2010). Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Molecular Ecology*, ***19***, 3565–3575.

Anderson, MJ, & Willis, TJ. (2003). Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. *Ecology*, ***84***, 511–525.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data.

Antunes, A, Troyer, JL, Roelke, ME, Pecon-Slattery, J, … Johnson, WE. (2008). The evolutionary dynamics of the lion *Panthera leo* revealed by host and viral population genomics. *PLoS Genetics*, ***4***.

Archie, EA, Luikart, G, & Ezenwa, VO. (2009). Infecting epidemiology with genetics: a new frontier in disease ecology. *Trends in Ecology and Evolution*, ***24***, 21–30.

Atwood, TC, Deliberto, TJ, Smith, HJ, Stevenson, JS, & Vercauteren, KC. (2009). Spatial ecology of raccoons related to cattle and bovine tuberculosis in northeastern Michigan. *Journal of Wildlife Management*, ***73***, 647–654.

Baele, G, Lemey, P, Bedford, T, Rambaut, A, … Alekseyenko, A V. (2012). Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Molecular Biology and Evolution*, ***29***, 2157–2167.

Balkenhol, N, Cushman, SA, Storfer, AT, & Waits, LP (Eds.). (2016a). *Landscape genetics: concepts, methods, applications*. Chichester, UK: Wiley Blackwell.

Balkenhol, N, Cushman, SA, Waits, LP, & Storfer, A. (2016b). Current status, future opportunities, and remaining challenges in landscape genetics. In N. Balkenhol, S. A. Cushman, A. T. Storfer, & L. P. Waits (Eds.), *Landscape Genetics: Concepts, Methods, Applications* (pp. 247–256). Chichester, UK: Wiley Blackwell.

Balkenhol, N, Gugerli, F, Cushman, SA, Waits, LP, … Wagner, HH. (2009a). Identifying future research needs in landscape genetics: where to from here? *Landscape Ecology*, ***24***, 455–463.

Balkenhol, N, Waits, LP, & Dezzani, RJ. (2009b). Statistical approaches in landscape genetics: an evaluation of methods for linking landscape and genetic data. *Ecography*, ***32***, 818–830.

Barton, HD, Gregory, AJ, Davis, R, Hanlon, CA, & Wisely, SM. (2010). Contrasting landscape epidemiology of two sympatric rabies virus strains. *Molecular Ecology*, ***19***,

2725–2738.

Bartoń, K. (2014). MuMIn: multi-model inference. R package version 1.10.0.

Becker, DJ, Streicker, DG, & Altizer, S. (2015). Linking anthropogenic resources to wildlife-pathogen dynamics: a review and meta-analysis. *Ecology Letters*, ***18***, 483–495.

Beier, P. (1993). Determining minimum habitat areas and habitat corridors for cougars. *Conservation Biology*, ***7***, 94–108.

Biek, R, Drummond, AJ, & Poss, M. (2006). A virus reveals population structure and recent demographic history of its carnivore host. *Science*, ***311***, 538–541.

Biek, R, Pybus, OG, Lloyd-Smith, JO, & Didelot, X. (2015). Measurably evolving pathogens in the genomic era. *Trends in Ecology and Evolution*, ***30***, 306–313.

Biek, R, & Real, LA. (2010). The landscape genetics of infectious disease emergence and spread. *Molecular Ecology*, ***19***, 3515–3531.

Biek, R, Rodrigo, AG, Holley, D, Drummond, AJ, … Poss, M. (2003). Epidemiology, genetic diversity, and evolution of endemic feline immunodeficiency virus in a population of wild cougars. *Journal of Virology*, ***77***, 9578–89.

Bierwagen, BG. (2007). Connectivity in urbanizing landscapes: the importance of habitat configuration, urban area size, and dispersal. *Urban Ecosystems*, ***10***, 29–42.

Blanchong, JA, Samuel, MD, Scribner, KT, Weckworth, B V, … Filcek, KB. (2008). Landscape genetics and the spatial distribution of chronic wasting disease. *Biology Letters*, ***4***, 130–133.

Boni, MF, Posada, D, & Feldman, MW. (2007). An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics*, ***176***, 1035–1047.

Bourhy, H, Nakouné, E, Hall, M, Nouvellet, P, … Rambaut, A. (2016). Revealing the micro-scale signature of endemic zoonotic disease transmission in an African urban setting. *PLoS Pathogens*, ***12***, 1–15.

Bowman, J, Greenhorn, JE, Marrotte, RR, McKay, MM, … Wehtje, M. (2016). On applications of landscape genetics. *Conservation Genetics*, ***17***, 753–760.

Bradley, CA, & Altizer, S. (2007). Urbanization and the ecology of wildlife diseases. *Trends in Ecology and Evolution*, ***22***, 95–102.

Brar, S, Tsui, CKM, Dhillon, B, Bergeron, M-J, … Hamelin, RC. (2015). Colonization history, host distribution, anthropogenic influence and landscape features shape populations of white pine blister rust, an invasive alien tree pathogen. *PLoS One*, ***10***, e0127916.

Brearley, G, Rhodes, J, Bradley, A, Baxter, G, … Mcalpine, C. (2013). Wildlife disease prevalence in human-modified landscapes. *Biological Reviews*, ***88***, 427–442.

Brown, JH, & Kodric-Brown, A. (1977). Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology*, ***58***, 445–449.

Brunker, K, Hampson, K, Horton, DL, & Biek, R. (2012). Integrating the landscape epidemiology and genetics of RNA viruses: rabies in domestic dogs as a model. *Parasitology*, ***139***, 1899–1913.

Brunker, K, Lemey, P, Marston, DA, Fooks, AR, … Biek, R. (2018). Landscape attributes governing local transmission of an endemic zoonosis: rabies virus in domestic dogs. *Molecular Ecology*, ***27***, 773–788.

Bull, CM, Godfrey, SS, & Gordon, DM. (2012). Social networks and the spread of Salmonella in a sleepy lizard population. *Molecular Ecology*, ***21***, 4386–4392.

Calonnec, A, Cartolaro, P, Naulin, JM, Bailey, D, & Langlais, M. (2008). A host-pathogen simulation model: powdery mildew of grapevine. *Plant Pathology*, ***57***, 493–508.

Caltrans. (2017). State Highway Network (SHN) and Postmile System.

Carrel, M, Patel, J, Taylor, SM, Janko, M, … Emch, M. (2015). The geography of malaria genetics in the Democratic Republic of Congo: a complex and fragmented landscape. *Social Science & Medicine*, ***133***, 233–241.

Carrel, MA, Emch, M, Tung, N, Jobe, RT, & Wan, X-F. (2012). Population-environment drivers of H5N1 avian influenza molecular change in Vietnam. *Health & Place*, ***18***, 1122–1131.

Carroll, MW, Matthews, DA, Hiscox, JA, Elmore, MJ, … Günther, S. (2015). Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa. *Nature*, ***524***, 97–101.

Carver, S, Bevins, SN, Lappin, MR, Boydston, EE, … Vandewoude, S. (2016). Pathogen exposure varies widely among sympatric populations of wild and domestic felids across the United States. *Ecological Applications*, ***26***, 367–381.

Catchen, J, Hohenlohe, PA, Bassham, S, Amores, A, & Cresko, WA. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, ***22***, 3124–3140.

Chang, X, Zhong, D, Lo, E, Fang, Q, … Yan, G. (2016). Landscape genetic structure and evolutionary genetics of insecticide resistance gene mutations in *Anopheles sinensis*. *Parasites & Vectors*, ***9***, 228.

Clarke, RT, Rothery, P, & Raybould, AF. (2002). Confidence limits for regression relationships between distance matrices: estimating gene flow with distance. *Journal of Agricultural, Biological, and Environmental Statistics*, ***7***, 361–372.

Corman, VM, Ithete, NL, Richards, LR, Schoeman, MC, … Drexler, JF. (2014). Rooting the phylogenetic tree of Middle East Respiratory Syndrome coronavirus by characterization of a conspecific virus from an African bat. *Journal of Virology*, ***88***, 11297–11303.

Cote, H, Garant, D, Robert, K, Mainguy, J, & Pelletier, F. (2012). Genetic structure and rabies spread potential in raccoons: the role of landscape barriers and sex-biased dispersal. *Evolutionary Applications*, ***5***, 393–404.

Craft, ME. (2015). Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society B: Biological Sciences*, ***370***, 20140107.

Craft, ME, & Caillaud, D. (2011). Network models: an underutilized tool in wildlife epidemiology? *Interdisciplinary Perspectives on Infectious Diseases*, ***2011***, 676949.

Craft, ME, Volz, E, Packer, C, & Meyers, LA. (2009). Distinguishing epidemic waves from disease spillover in a wildlife population. *Proceedings of the Royal Society B: Biological*

Sciences, ***276***, 1777–1785.

Crooks, KR. (2002). Relative sensitivites of mammalian carnivores to habitat fragmentation. *Conservation Biology*, ***16***, 488–502.

Crooks, KR, Riley, SPD, Gehrt, SD, Gosselink, TE, & van Deelen, TR. (2010). Community Ecology of Urban Carnivores. In S. D. Gehrt, S. P. D. Riley, & B. L. Cypher (Eds.), *Urban Carnivores: Ecology, Conflict, and Conservation* (pp. 185–198). Baltimore, Maryland, USA: The Johns Hopkins University Press.

Crooks, KR, & Soulé, ME. (1999). Mesopredator release and avifaunal extinctions in a fragmented system. *Nature*, ***400***, 563–566.

Cullingham, CI, Kyle, CJ, Pond, BA, Rees, EE, & White, BN. (2009). Differential permeability of rivers to raccoon gene flow corresponds to rabies incidence in Ontario, Canada. *Molecular Ecology*, ***18***, 43–53.

Cushman, SA, & Landguth, EL. (2010a). Scale dependent inference in landscape genetics. *Landscape Ecology*, ***25***, 967–979.

Cushman, SA, & Landguth, EL. (2010b). Spurious correlations and inference in landscape genetics. *Molecular Ecology*, ***19***, 3592–3602.

Cushman, SA, Wasserman, TN, Landguth, EL, & Shirk, AJ. (2013). Re-evaluating causal modeling with Mantel tests in landscape genetics. *Diversity*, ***5***, 51–72.

de Wit, P, & Palumbi, SR. (2013). Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. *Molecular Ecology*, ***22***, 2884–2897.

DeCandia, AL, Dobson, AP, & vonHoldt, BM. (2018). Toward an integrative molecular approach to wildlife disease. *Conservation Biology*, ***32***, 798–807.

Decout, S, Manel, S, Miaud, C, & Luque, S. (2012). Integrative approach for landscape-based graph connectivity analysis: a case study with the common frog (*Rana temporaria*) in human-dominated landscapes. *Landscape Ecology*, ***27***, 267–279.

Dellicour, S, Rose, R, Faria, NR, Lemey, P, & Pybus, OG. (2016a). SERAPHIM: Studying environmental rasters and phylogenetically informed movements. *Bioinformatics*, ***32***, 3204–3206.

Dellicour, S, Rose, R, Faria, NR, Vieira, LFP, … Pybus, OG. (2017). Using viral gene sequences to compare and explain the heterogeneous spatial dynamics of virus epidemics. *Molecular Biology and Evolution*, ***34***, 2563–2571.

Dellicour, S, Rose, R, & Pybus, OG. (2016b). Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinformatics*, ***17***, 82.

DeYoung, RW, Zamorano, A, Mesenbrink, BT, Campbell, TA, … Root, JJ. (2009). Landscape-genetic analysis of population structure in the Texas gray fox oral rabies vaccination zone. *Journal of Wildlife Management*, ***73***, 1292–1299.

Didelot, X, Fraser, C, Gardy, J, Colijn, C, & Malik, H. (2017). Genomic infectious disease epidemiology in partially sampled and ongoing outbreaks. *Molecular Biology and Evolution*, ***34***, 997–1007.

Dixon, JD, Oli, MK, Wooten, MC, Eason, TH, … Cunningham, MW. (2007). Genetic consequences of habitat fragmentation and loss: the case of the Florida black bear (*Ursus americanus floridanus*). *Conservation Genetics*, ***8***, 455–464.

Do, C, Waples, RS, Peel, D, Macbeth, GM, … Ovenden, JR. (2014). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, ***14***, 209–214.

Dobson, AP, Rodriguez, JP, Roberts, WM, & Wilcove, DS. (1997). Geographic distribution of endangered species in the United States. *Science*, ***275***, 550–553.

Dougherty, ER, Seidel, DP, Carlson, CJ, Spiegel, O, & Getz, WM. (2018). Going through the motions: incorporating movement analyses into disease research. *Ecology Letters*, ***21***, 588–604.

Drummond, AJ, Nicholls, GK, Rodrigo, AG, & Solomon, W. (2002). Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*, ***161***, 1307–1320.

Drummond, AJ, Suchard, MA, Xie, D, & Rambaut, A. (2012). Bayesian Phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, ***29***, 1969–1973.

Duforet-Frebourg, N, & Blum, MGB. (2014). Nonstationary patterns of isolation-by-distance: inferring measures of local genetic differentiation with bayesian kriging. *Evolution*, ***68***, 1110–1123.

Dyer, RJ. (2015). Is there such a thing as landscape genetics? *Molecular Ecology*, ***24***, 3518–3528.

Dyer, RJ, & Nason, JD. (2004). Population Graphs: the graph theoretic shape of genetic structure. *Molecular Ecology*, ***13***, 1713–1727.

Dyer, RJ, Nason, JD, & Garrick, RC. (2010). Landscape modelling of gene flow: improved power using conditional genetic distance derived from the topology of population networks. *Molecular Ecology*, ***19***, 3746–3759.

Ellis, AM, Václavík, T, & Meentemeyer, RK. (2010). When is connectivity important? A case study of the spatial pattern of sudden oak death. *Oikos*, ***119***, 485–493.

Elveback, L, & Varma, A. (1965). Simulation of mathematical models for public health problems. *Public Health Reports*, ***80***, 1067–1076.

Epstein, B, Jones, M, Hamede, R, Hendricks, S, … Storfer, A. (2016). Rapid evolutionary response to a transmissible cancer in Tasmanian devils. *Nature Communications*, ***7***, 12684.

Ernest, HB, Vickers, TW, Morrison, SA, Buchalski, MR, & Boyce, WM. (2014). Fractured genetic connectivity threatens a Southern California puma (*Puma concolor*) population. *PLoS ONE*, ***9***.

Escobar, LE, Carver, S, Romero-Alvarez, D, VandeWoude, S, … Craft, ME. (2017). Inferring the ecological niche of Toxoplasma gondii and Bartonella spp. in wild felids. *Frontiers in Veterinary Science*, ***4***, 172.

Estes, JA, Terborgh, J, Brashares, JS, Power, ME, … Wardle, DA. (2011). Trophic downgrading of planet Earth. *Science*, ***333***, 301–306.

Evans, JS, Oakleaf, J, Cushman, SA, & Theobald, D. (2014). An ArcGIS toolbox for surface gradient and geomorphometric modeling, version 2.0-0. Laramie, WY.

Excoffier, L, Hofer, T, & Foll, M. (2009). Detecting loci under selection in a hierarchically structured population. *Heredity*, ***103***, 285–298.

Fahrig, L. (2003). Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, ***34***, 487–515.

Fahrig, L, Baudry, J, Brotons, L, Burel, FG, … Martin, J-L. (2011). Functional landscape heterogeneity and animal biodiversity in agricultural landscapes. *Ecology Letters*, ***14***, 101–112.

Faria, NR, Rambaut, A, Suchard, MA, Baele, G, … Lemey, P. (2014). The early spread and epidemic ignition of HIV-1 in human populations. *Science*, ***346***, 56–61.

Faria, NR, Suchard, MA, Rambaut, A, Streicker, DG, & Lemey, P. (2013). Simultaneously reconstructing viral cross-species transmission history and identifying the underlying constraints. *Philosophical Transactions of the Royal Society B: Biological Sciences*, ***368***, 20120196.

Faust, CL, McCallum, HI, Bloomfield, LSP, Gottdenker, NL, … Plowright, RK. (2018). Pathogen spillover during land conversion. *Ecology Letters*, ***21***, 471–483.

Ferrier, S, & Guisan, A. (2006). Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, ***43***, 393–404.

Fischer, MC, Rellstab, C, Leuzinger, M, Roumet, M, … Widmer, A. (2017). Estimating genomic diversity and population differentiation - an empirical comparison of microsatellite and SNP variation in Arabidopsis halleri. *BMC Genomics*, ***18***, 1–15.

Fitzpatrick, MC, & Keller, SR. (2015). Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, ***18***, 1–16.

Fountain-Jones, NM, Craft, ME, Funk, WC, Kozakiewicz, C, … Carver, S. (2017a). Urban landscapes can change virus gene flow and evolution in a fragmentation-sensitive carnivore. *Molecular Ecology*, ***26***, 6487–6498.

Fountain-Jones, NM, Packer, C, Troyer, JL, VanderWaal, K, … Craft, ME. (2017b). Linking social and spatial networks to viral community phylogenetics reveals subtype-specific transmission dynamics in African lions. *Journal of Animal Ecology*, ***86***, 1469–1482.

Fountain-Jones, NM, Pearse, WD, Escobar, LE, Alba-Casals, A, … Craft, ME. (2017c). Towards an eco-phylogenetic framework for infectious disease ecology. *Biological Reviews*, ***93***, 950–970.

François, O, & Waits, LP. (2016). Clustering and assignment methods in landscape genetics. In N. Balkenhol, S. A. Cushman, A. T. Storfer, & L. P. Waits (Eds.), *Landscape Genetics: Concepts, Methods, Applications* (pp. 114–128). Chichester, UK: Wiley Blackwell.

Frantz, AC, Cellina, S, Krier, A, Schley, L, & Burke, T. (2009). Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: clusters or isolation by distance? *Journal of Applied Ecology*, ***46***, 493–505.

Fraser, D, Mouton, A, Serieys, LEK, Cole, S, … Wayne, R. (2018). Genome-wide expression reveals multiple systemic effects associated with detection of anticoagulant poisons in bobcats (*Lynx rufus*). *Molecular Ecology*, ***27***, 1170–1187.

Frichot, E, & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, ***6***, 925–929.

Frichot, E, Mathieu, F, Trouillon, T, Bouchard, G, & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, ***196***, 973–983.

Fuller, TK, DeStefano, S, & Warren, PS. (2010). Carnivore behavior and ecology, and relationship to urbanization. In *Urban Carnivores: Ecology, Conflict, and Conservation* (pp. 13–20).

Fumagalli, M, Sironi, M, Pozzoli, U, Ferrer-Admettla, A, … Nielsen, R. (2011). Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genetics*, ***7***, e1002355.

Funk, WC, Blouin, MS, Corn, PS, Maxell, BA, … Allendorf, FW. (2005). Population structure of Columbia spotted frogs (*Rana luteiventris*) is strongly affected by the landscape. *Molecular Ecology*, ***14***, 483–496.

Garroway, CJ, Bowman, J, Carr, D, & Wilson, PJ. (2008). Applications of graph theory to landscape genetics. *Evolutionary Applications*, ***1***, 620–630.

Garroway, CJ, Radersma, R, Sepil, I, Santure, AW, … Sheldon, BC. (2013). Fine-scale genetic structure in a wild bird population: the role of limited dispersal and environmentally based selection as causal factors. *Evolution*, ***67***, 3488–3500.

Gibbs, MJ, Armstrong, JS, & Gibbs, AJ. (2000). Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics*, ***16***, 573–582.

Gilbertson, MLJ, Fountain-Jones, NM, & Craft, ME. (2018). Incorporating genomic methods into contact networks to reveal new insights into animal behaviour and infectious disease dynamics. *Behaviour*, ***155***, 759–791.

Gillies, CS, & St. Clair, CC. (2008). Riparian corridors enhance movement of a forest specialist bird in fragmented tropical forest. *Proceedings of the National Academy of Sciences*, ***105***, 19774–19779.

Goldberg, CS, & Waits, LP. (2010). Comparative landscape genetics of two pond-breeding amphibian species in a highly modified agricultural landscape. *Molecular Ecology*, ***19***, 3650–3663.

Gordon, ER, Curns, AT, Krebs, JW, Rupprecht, CE, … Childs, JE. (2004). Temporal dynamics of rabies in a wildlife host and the risk of cross-species transmission. *Epidemiology and Infection*, ***132***, 515–524.

Gottdenker, NL, Streicker, DG, Faust, CL, & Carroll, CR. (2014). Anthropogenic land use change and infectious diseases: a review of the evidence. *EcoHealth*, ***11***, 619–632.

Grenfell, BT, Pybus, OG, Gog, JR, Wood, JR, … Holmes, EC. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*, ***303***, 327–332.

Grover, CE, Salmon, A, & Wendel, JF. (2012). Targeted sequence capture as a powerful tool for evolutionary analysis. *American Journal of Botany*, ***99***, 312–319.

Guillot, G, Mortier, F, & Estoup, A. (2005). GENELAND: A computer package for landscape genetics. *Molecular Ecology Notes*, ***5***, 712–715.

Hamede, RK, Bashford, J, McCallum, H, & Jones, M. (2009). Contact networks in a wild Tasmanian devil (*Sarcophilus harrisii*) population: using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease. *Ecology Letters*, ***12***, 1147–1157.

Harmon, LJ, Weir, JT, Brock, CD, Glor, RE, & Challenger, W. (2008). GEIGER: Investigating evolutionary radiations. *Bioinformatics*, ***24***, 129–131.

Helyar, SJ, Hemmer-Hansen, J, Bekkevold, D, Taylor, MI, … Nielsen, EE. (2011). Application of SNPs for population genetics of nonmodel organisms: New opportunities and challenges. *Molecular Ecology Resources*, ***11***, 123–136.

Hilty, JA, & Merenlender, AM. (2004). Use of riparian corridors and vineyards by mammalian predators in northern California. *Conservation Biology*, ***18***, 126–135.

Hoffmann, AA, Sgrò, CM, & Kristensen, TN. (2017). Revisiting adaptive potential, population size, and conservation. *Trends in Ecology and Evolution*, ***32***, 506–517.

Holderegger, R, & Wagner, HH. (2008). Landscape genetics. *BioScience*, ***58***, 199–207.

Holmala, K, & Kauhala, K. (2006). Ecology of wildlife rabies in Europe. *Mammal Review*, ***36***, 17–36.

Homer, C, Dewitz, J, Yang, L, Jin, S, … Megown, K. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States – representing a decade of land cover change information. *Photogrammetric Engineering & Remote Sensing*, ***81***, 345–354.

Jacquot, M, Nomikou, K, Palmarini, M, Mertens, P, & Biek, R. (2017). Bluetongue virus spread in Europe is a consequence of climatic, landscape and vertebrate host factors as revealed by phylogeographic inference. *Proceedings of the Royal Society B: Biological Sciences*, ***284***, 20170919.

Jay, F. (2011). POPS: Prediction of population genetic structure. Program documentation, University Joseph Fourier Grenoble, France.

Jay, F, Manel, S, Alvarez, N, Durand, EY, … François, O. (2012). Forecasting changes in population genetic structure of alpine plants in response to global warming. *Molecular Ecology*, ***21***, 2354–2368.

Jennings, M, & Lewison, R. (2013). Planning for connectivity under climate change: using bobcat movement to assess landscape connectivity across San Diego County's open spaces. *Report for the San Diego Foundation, Blasker Final Report*, 50 pp.

Jennings, M, & Zeller, K. (2017). Comprehensive multi-species connectivity assessment and planning for the highway 67 region of San Diego County, California. *Final Report for the San Diego Association of Governments*, 77 pp.

Joannon, B, Lavigne, C, Lecoq, H, & Desbiez, C. (2010). Barriers to gene flow between emerging populations of watermelon mosaic virus in southeastern France. *Phytopathology*, ***100***, 1373–1379.

Johnson, PTJ, de Roode, JC, & Fenton, A. (2015). Why infectious disease research needs

community ecology. *Science*, ***349***, 1259504.

Jombart, T, Devillard, S, Dufour, AB, & Pontier, D. (2008). Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, ***101***, 92–103.

Kalinowski, ST. (2005). HP-RARE 1.0: A computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, ***5***, 187–189.

Kaltz, O, & Shykoff, JA. (1998). Local adaptation in host – parasite systems. *Heredity*, ***81***, 361–370.

Kao, RR, Haydon, DT, Lycett, SJ, & Murcia, PR. (2014). Supersize me: how whole-genome sequencing and big data are transforming epidemiology. *Trends in Microbiology*, ***22***, 282–291.

Kass, RE, & Raftery, AE. (1995). Bayes Factors. *Journal of the American Statistical Association*, ***90***, 773–795.

Keenan, K, Mcginnity, P, Cross, TF, Crozier, WW, & Prodöhl, PA. (2013). DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, ***4***, 782–788.

Kelly, AC, Mateus-Pinilla, NE, Brown, W, Ruiz, MO, … Novakofski, J. (2014). Genetic assessment of environmental features that influence deer dispersal: implications for prion-infected populations. *Population Ecology*, ***56***, 327–340.

Kerr, TJ, Matthee, S, Govender, D, Tromp, G, … Matthee, CA. (2018). Viruses as indicators of contemporary host dispersal and phylogeography: an example of feline immunodeficiency virus (FIV_{Ple}) in free-ranging African lion (*Panthera leo*). *Journal of Evolutionary Biology*, ***31***, 1529–1543.

Keyghobadi, N. (2007). The genetic implications of habitat fragmentation for animals. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, ***85***, 1049–1064.

Kozakiewicz, CP, Burridge, CP, & Carver, S. (*in prep*) Does the virus cross the road? Viral phylogeographic patterns among urban bobcat populations.

Kozakiewicz, CP, Burridge, CP, & Carver, S. *(in prep)* Virus phylogeography in urban bobcat populations reveals increased transmission in vegetated habitats.

Kozakiewicz, CP, Burridge, CP, Funk, WC, Salerno, PE, … Carver, S. *(in review)* Urbanization reduces genetic connectivity in bobcats *(Lynx rufus)* at both intra- and inter-population spatial scales.

Kozakiewicz, CP, Burridge, CP, Funk, WC, Vandewoude, S, … Carver, S. (2018a). Pathogens in space: advancing understanding of pathogen dynamics and disease ecology through landscape genetics. *Evolutionary Applications*, ***11***, 1763–1778.

Kozakiewicz, CP, Carver, S, & Burridge, CP. (2018b). Under-representation of avian studies in landscape genetics. *Ibis*, ***160***, 1–12.

Lagana, DM, Lee, JS, Lewis, JS, Bevins, SN, … VandeWoude, S. (2013). Characterization of regionally associated feline immunodeficiency virus (FIV) in bobcats (*Lynx rufus*). *Journal of Wildlife Diseases*, ***49***, 718–722.

Landguth, E, Cushman, SA, & Balkenhol, N. (2016). Simulation modeling in landscape genetics. In N. Balkenhol, S. A. Cushman, A. T. Storfer, & L. P. Waits (Eds.),

Landscape Genetics: Concepts, Methods, Applications (pp. 101–113). Chichester, UK: Wiley Blackwell.

Landguth, EL, Cushman, SA, Schwartz, MK, McKelvey, KS, … Luikart, G. (2010). Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology*, ***19***, 4179–4191.

Landguth, EL, Holden, ZA, Mahalovich, MF, & Cushman, SA. (2017). Using landscape genetics simulations for planting blister rust resistant whitebark pine in the US Northern Rocky Mountains. *Frontiers in Genetics*, ***8***, 1–12.

Larson, WA, Seeb, JE, Dann, TH, Schindler, DE, & Seeb, LW. (2014). Signals of heterogeneous selection at an MHC locus in geographically proximate ecotypes of sockeye salmon. *Molecular Ecology*, ***23***, 5448–5461.

LaRue, MA, & Nielsen, CK. (2008). Modelling potential dispersal corridors for cougars in midwestern North America using least-cost path methods. *Ecological Modelling*, ***212***, 372–381.

Lavery, J V, Harrington, LC, & Scott, TW. (2008). Ethical, social, and cultural considerations for site selection for research with genetically modified mosquitoes. *The American Journal of Tropical Medicine and Hygiene*, ***79***, 312–318.

Lee, JS, Bevins, SN, Serieys, LE, Vickers, W, … VandeWoude, S. (2014). Evolution of puma lentivirus in bobcats (*Lynx rufus*) and mountain lions (*Puma concolor*) in North America. *Journal of Virology*, ***88***, 7727–7737.

Lee, JS, Mackie, RS, Harrison, T, Shariat, B, … VandeWoude, S. (2017). Targeted enrichment for pathogen detection and characterization in three felid species. *Journal of Clinical Microbiology*, ***55***, 1658–1670.

Lee, JS, Ruell, EW, Boydston, EE, Lyren, LM, … Vandewoude, S. (2012). Gene flow and pathogen transmission among bobcats (*Lynx rufus*) in a fragmented urban landscape. *Molecular Ecology*, ***21***, 1617–1631.

Lefèvre, T, Adamo, SA, Biron, DG, Missé, D, … Thomas, F. (2009). Chapter 3. Invasion of the body snatchers: the diversity and evolution of manipulative strategies in host-parasite interactions. *Advances in Parasitology*, ***68***, 45–83.

Lemaire, C, De Gracia, M, Leroy, T, Michalecka, M, … Le Cam, B. (2016). Emergence of new virulent populations of apple scab from nonagricultural disease reservoirs. *New Phytologist*, ***209***, 1220–1229.

Lemey, P, Rambaut, A, Welch, JJ, & Suchard, MA. (2010). Phylogeography takes a relaxed random walk in continuous space and time. *Molecular Biology and Evolution*, ***27***, 1877–1885.

Leo, SST, Gonzalez, A, Millien, V, & Cristescu, ME. (2016). Multi-taxa integrated landscape genetics for zoonotic infectious diseases: deciphering variables influencing disease emergence. *Genome*, ***59***, 349–361.

Leroy, T, Lemaire, C, Dunemann, F, & Le Cam, B. (2013). The genetic structure of a *Venturia inaequalis* population in a heterogeneous host population composed of different *Malus* species. *BMC Evolutionary Biology*, ***13***.

Leu, ST, Kappeler, PM, & Bull, CM. (2010). Refuge sharing network predicts ectoparasite

load in a lizard. *Behavioral Ecology and Sociobiology*, ***64***, 1495–1503.

Liang, L, Liu, Y, Liao, J, & Gong, P. (2014). Wetlands explain most in the genetic divergence pattern of *Oncomelania hupensis*. *Infection Genetics and Evolution*, ***27***, 436–444.

Lion, S, & Gandon, S. (2015). Evolution of spatially structured host-parasite interactions. *Journal of Evolutionary Biology*, ***28***, 10–28.

Lo, E, Hemming-Schroeder, E, Yewhalaw, D, Nguyen, J, … Yan, G. (2017). Transmission dynamics of co-endemic *Plasmodium vivax* and *P. falciparum* in Ethiopia and prevalence of antimalarial resistant genotypes. *PLoS Neglected Tropical Diseases*, ***11***, 1–25.

Lotterhos, KE, & Whitlock, MC. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, ***24***, 1031–1046.

Lu, Z, Johnson, WE, Menotti-Raymond, M, Yuhki, N, … O'Brien, SJ. (2001). Patterns of genetic diversity in remaining giant panda populations. *Conservation Biology*, ***15***, 1596–1607.

Luu, K, Bazin, E, & Blum, MGB. (2017). *pcadapt*: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, ***17***, 67–77.

Lyren, LM, Alonso, RS, Crooks, KR, & Boydston, EE. (2008a). Evaluation of functional connectivity for bobcats and coyotes across the former El Toro Marine Base, Orange County, California. *Administrative Report*, 179.

Lyren, LM, Alonso, RS, Crooks, KR, & Boydston, EE. (2008b). GPS telemetry, camera trap, and mortality surveys of bobcats in the San Joaquin Hills, Orange County, California. *Administrative Report*, 134 pp.

Lyren, LM, Turschak, GM, Ambat, ES, Haas, CD, … Crooks, KR. (2006). Carnivore Activity and Movement in a Southern California Protected Area, the North/Central Irvine Ranch. *U.S. Geological Survey Technical Report*, 115 pp.

Machtans, CS, Villard, MA, & Hannon, SJ. (1996). Use of riparian buffer strips as movement corridors by forest birds. *Conservation Biology*, ***10***, 1366–1379.

Maddison, WP. (1997). Gene trees in species trees. *Systematic Biology*, ***46***, 523–536.

Magee, D, Beard, R, Suchard, MA, Lemey, P, & Scotch, M. (2015). Combining phylogeography and spatial epidemiology to uncover predictors of H5N1 influenza A virus diffusion. *Archives of Virology*, ***160***, 215–224.

Manel, S, & Holderegger, R. (2013). Ten years of landscape genetics. *Trends in Ecology and Evolution*, ***28***, 614–621.

Manel, S, Schwartz, MK, Luikart, G, & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, ***18***, 189–197.

Markovchick-Nicholls, L, Regan, HM, Deutschman, DH, Widyanata, A, … Ann Hunt, T. (2008). Relationships between human disturbance and wildlife land use in urban habitat

fragments. *Conservation Biology*, ***22***, 99–109.

Martin, DP, Murrell, B, Golden, M, Khoosal, A, & Muhire, B. (2015). RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evolution*, ***1***, 1–5.

Martin, DP, & Rybicki, E. (2000). RDP: detection of recombination amongst aligned sequences. *Bioinformatics*, ***16***, 562–563.

Mastretta-Yanes, A, Arrigo, N, Alvarez, N, Jorgensen, TH, … Emerson, BC. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources*, ***15***, 28–41.

Maynard Smith, J. (1992). Analyzing the mosaic structure of genes. *Journal of Molecular Evolution*, ***34***, 126–129.

Mazé-Guilmo, E, Blanchet, S, Mccoy, KD, & Loot, G. (2016). Host dispersal as the driver of parasite genetic structure: A paradigm lost? *Ecology Letters*, ***19***, 336–347.

McCallum, H, & Dobson, A. (2002). Disease, habitat fragmentation and conservation. *Proceedings of the Royal Society B: Biological Sciences*, ***269***, 2041–2049.

McRae, BH. (2006). Isolation by resistance. *Evolution*, ***60***, 1551–1561.

McRae, BH, Dickson, BG, Keitt, TH, & Shah, VB. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, ***89***, 2712–2724.

Meentemeyer, RK, Haas, SE, & Václavík, T. (2012). Landscape epidemiology of emerging infectious diseases in natural and human-altered ecosystems. *Annual Review of Phytopathology*, ***50***, 379–402.

Monello, RJ, Galloway, NL, Powers, JG, Madsen-Bouterse, SA, … Wild, MA. (2017). Pathogen-mediated selection in free-ranging elk populations infected by chronic wasting disease. *Proceedings of the National Academy of Sciences*, ***114***, 12208–12212.

Monmonier, MS. (1973). Maximum-difference barriers: an alternative numerical regionalization method. *Geographical Analysis*, ***5***, 245–261.

Muhire, BM, Varsani, A, & Martin, DP. (2014). SDT: A virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS ONE*, ***9***.

Murcia, C. (1995). Edge effects in fragmented forests: implications for conservation. *Trends in Ecology & Evolution*, ***10***, 58–62.

Murphy, MA, Dezzani, R, Pilliod, DS, & Storfer, A. (2010a). Landscape genetics of high mountain frog metapopulations. *Molecular Ecology*, ***19***, 3634–3649.

Murphy, MA, Evans, JS, & Storfer, A. (2010b). Quantifying *Bufo boreas* connectivity in Yellowstone National Park with landscape genetics. *Ecology*, ***91***, 252–261.

Murrell, B, Weaver, S, Smith, MD, Wertheim, JO, … Kosakovsky Pond, SL. (2015). Gene-wide identification of episodic selection. *Molecular Biology and Evolution*, ***32***, 1365–1371.

Myers, N, Mittermeier, RA, Mittermeier, CG, da Fonseca, GAB, & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, ***403***, 853–858.

Nelson, HJ. (1959). The spread of an artificial landscape over southern California. *Annals of*

the Association of American Geographers, ***49***, 80–99.

Newman, MEJ. (2003). The structure and function of complex networks. *SIAM Review*, ***45***, 167–256.

Noss, RF, Quigley, HB, Hornocker, MG, Merrill, T, & Paquet, PC. (1996). Conservation biology and carnivore conservation in the Rocky Mountains. *Conservation Biology*, ***10***, 949–963.

Ordeñana, MA, Crooks, KR, Boydston, EE, Fisher, RN, … Van Vuren, DH. (2010). Effects of urbanization on carnivore species distribution and richness. *Journal of Mammalogy*, ***91***, 1322–1331.

Ostfeld, RS, Glass, GE, & Keesing, F. (2005). Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in Ecology and Evolution*, ***20***, 328–336.

Padidam, M, Sawyer, S, & Fauquet, CM. (1999). Possible emergence of new geminiviruses by frequent recombination. *Virology*, ***265***, 218–225.

Pamilo, P, & Nei, M. (1988). Relationships between gene trees and species trees. *Molecular Biology and Evolution*, ***5***, 568–583.

Paris, JR, Stevens, JR, & Catchen, JM. (2017). Lost in parameter space: a road map for Stacks. *Methods in Ecology and Evolution*, ***8***, 1360–1373.

Pease, JB, & Hahn, MW. (2013). More accurate phylogenies inferred from low-recombination regions in the presence of incomplete lineage sorting. *Evolution*, ***67***, 2376–2384.

Pease, KM, Freedman, AH, Pollinger, JP, McCormack, JE, … Wayne, RK. (2009). Landscape genetics of California mule deer (*Odocoileus hemionus*): the roles of ecological and historical factors in generating differentiation. *Molecular Ecology*, ***18***, 1848–1862.

Perkins, SE, Cagnacci, F, Stradiotto, A, Arnoldi, D, & Hudson, PJ. (2009). Comparison of social networks derived from ecological data: implications for inferring infectious disease dynamics. *Journal of Animal Ecology*, ***78***, 1015–1022.

Peterson, BK, Weber, JN, Kay, EH, Fisher, HS, & Hoekstra, HE. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, ***7***.

Petkova, D, Novembre, J, & Stephens, M. (2016). Visualizing spatial population structure with estimated effecitive migration surfaces. *Nature Genetics*, ***48***, 94–100.

Pickett, STA, & Cadenasso, ML. (1995). Landscape ecology: spatial heterogeneity in ecological systems. *Science*, ***269***, 331–334.

Piry, S, Chapuis, MP, Gauffre, B, Papaïx, J, … Berthier, K. (2016). Mapping Averaged Pairwise Information (MAPI): a new exploratory tool to uncover spatial structure. *Methods in Ecology and Evolution*, ***7***, 1463–1475.

Poessel, SA, Burdett, CL, Boydston, EE, Lyren, LM, … Crooks, KR. (2014). Roads influence movement and home ranges of a fragmentation-sensitive carnivore, the bobcat, in an urban landscape. *Biological Conservation*, ***180***, 224–232.

Pond, SLK, & Muse, S V. (2005). HyPhy: hypothesis testing using phylogenies. In *Statistical*

Methods in Molecular Evolution (pp. 125–181). New York, NY: Springer.

Posada, D, & Crandall, KA. (2001). Evaluation of methods for detecting recombination from DNA sequences: computer simulations. *Proceedings of the National Academy of Sciences*, ***98***, 13757–13762.

Prentice, JC, Marion, G, White, PCL, Davidson, RS, & Hutchings, MR. (2014). Demographic processes drive increases in wildlife disease following population reduction. *PLoS ONE*, ***9***, e86563.

Pritchard, JK, Stephens, M, & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, ***155***, 945–959.

Purcell, S, Neale, B, Todd-Brown, K, Thomas, L, … Sham, PC. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, ***81***, 559–575.

R Development Core Team. (2013). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*.

Raj, A, Stephens, M, & Pritchard, JK. (2014). FastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics*, ***197***, 573–589.

Ramalho, CE, & Hobbs, RJ. (2012). Time for a change: dynamic urban ecology. *Trends in Ecology and Evolution*, ***27***, 179–188.

Rambaut, A, Drummond, AJ, Xie, D, Baele, G, & Suchard, MA. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, ***67***, 901–904.

Randa, LA, & Yunger, JA. (2006). Carnivore occurrence along an urban–rural gradient: a landscape-level analysis. *Journal of Mammalogy*, ***87***, 1154–1164.

Real, L, & Biek, R. (2007). Spatial dynamics and genetics of infectious diseases on heterogeneous landscapes. *Journal of the Royal Society Interface*, ***4***, 935–948.

Real, LA, Henderson, JC, Biek, R, Snaman, J, … Nadin-Davis, S. (2005). Unifying the spatial population dynamics and molecular evolution of epidemic rabies virus. *Proceedings of the National Academy of Sciences*, ***102***, 12107–12111.

Reding, DM, Bronikowski, AM, Johnson, WE, & Clark, WR. (2012). Pleistocene and ecological effects on continental-scale genetic differentiation in the bobcat (*Lynx rufus*). *Molecular Ecology*, ***21***, 3078–93.

Rees, EE, Pond, BA, Cullingham, CI, Tinline, R, … White, BN. (2008). Assessing a landscape barrier using genetic simulation modelling: implications for raccoon rabies management. *Preventive Veterinary Medicine*, ***86***, 107–123.

Rellstab, C, Gugerli, F, Eckert, AJ, Hancock, AM, & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, ***24***, 4348–4370.

Richardson, JL, Brady, SP, Wang, IJ, & Spear, SF. (2016). Navigating the pitfalls and promise of landscape genetics. *Molecular Ecology*, ***25***, 849–863.

Rieux, A, De Bellaire, LDL, Zapater, MF, Ravigne, V, & Carlier, J. (2013). Recent range expansion and agricultural landscape heterogeneity have only minimal effect on the

spatial genetic structure of the plant pathogenic fungus *Mycosphaerella fijiensis*. *Heredity*, ***110***, 29–38.

Rieux, A, Halkett, F, de Bellaire, L de L, Zapater, MF, … Carlier, J. (2011). Inferences on pathogenic fungus population structures from microsatellite data: new insights from spatial genetics approaches. *Molecular Ecology*, ***20***, 1661–1674.

Riley, SPD. (2006). Spatial ecology of bobcats and gray foxes in urban and rural zones of a national park. *The Journal of Wildlife Management*, ***70***, 1425–1435.

Riley, SPD, Boydston, EE, Crooks, KR, & Lyren, LM. (2010). Bobcats (Lynx rufus). In S. D. Gehrt, S. P. D. Riley, & B. L. Cypher (Eds.), *Urban Carnivores: Ecology, Conflict, and Conservation* (pp. 121–138). Baltimore, Maryland, USA: The Johns Hopkins University Press.

Riley, SPD, Bromley, C, Poppenga, RH, Uzal, FA, … Sauvajot, RM. (2007). Anticoagulant exposure and notoedric mange in bobcats and mountain lions in urban southern California. *Journal of Wildlife Management*, ***71***, 1874–1884.

Riley, SPD, Foley, J, & Chomel, B. (2004). Exposure to feline and canine pathogens in bobcats and gray foxes in urban and rural zones of a national park in California. *Journal of Wildlife Diseases*, ***40***, 11–22.

Riley, SPD, Pollinger, JP, Sauvajot, RM, York, EC, … Wayne, RK. (2006). A southern California freeway is a physical and social barrier to gene flow in carnivores. *Molecular Ecology*, ***15***, 1733–1741.

Riley, SPD, Sauvajot, RM, Fuller, TK, York, EC, … Wayne, RK. (2003). Effects of urbanization and habitat fragmentation on bobcats and coyotes in southern California. *Conservation Biology*, ***17***, 566–576.

Riley, SPD, Serieys, LEK, Pollinger, JP, Sikich, JA, … Ernest, HB. (2014). Individual behaviors dominate the dynamics of an urban mountain lion population isolated by roads. *Current Biology*, ***24***, 1989–1994.

Rioux Paquette, S, Talbot, B, Garant, D, Mainguy, J, & Pelletier, F. (2014). Modelling the dispersal of the two main hosts of the raccoon rabies variant in heterogeneous environments with landscape genetics. *Evolutionary Applications*, ***7***, 734–749.

Robertson, JM, Murphy, MA, Pearl, CA, Adams, MJ, … Funk, WC. (2018). Regional variation in drivers of connectivity for two frog species (*Rana pretiosa* and *R. luteiventris*) from the U.S. Pacific Northwest. *Molecular Ecology*, ***27***, 3242–3256.

Robinson, SJ, Samuel, MD, Rolley, RE, & Shelton, P. (2013). Using landscape epidemiological models to understand the distribution of chronic wasting disease in the Midwestern USA. *Landscape Ecology*, ***28***, 1923–1935.

Roelke, ME, Martenson, JS, & O'Brien, SJ. (1993). The consequences of demographic reduction and genetic depletion in the endangered Florida panther. *Current Biology*, ***3***, 340–350.

Roffler, GH, Amish, SJ, Smith, S, Cosart, T, … Luikart, G. (2016). SNP discovery in candidate adaptive genes using exon capture in a free-ranging alpine ungulate. *Molecular Ecology Resources*, ***16***, 1147–1164.

Row, JR, Knick, ST, Oyler-McCance, SJ, Lougheed, SC, & Fedy, BC. (2017). Developing

approaches for linear mixed modeling in landscape genetics through landscape-directed dispersal simulations. *Ecology and Evolution*, ***7***, 3751–3761.

Rozas, J, Ferrer-Mata, A, Sánchez-DelBarrio, JC, Guirao-Rico, S, … Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution*, ***34***, 3299–3302.

Ruell, EW, Riley, SPD, Douglas, MR, Antolin, MF, … Crooks, KR. (2012). Urban habitat fragmentation and genetic population structure of bobcats in coastal southern California. *The American Midland Naturalist*, ***168***, 265–280.

Salminen, MO, Carr, JK, Burke, DS, & McCutchan, FE. (1995). Identification of breakpoints in intergenotypic recombinants of HIV Type 1 by bootscanning. *AIDS Research and Human Retroviruses*, ***11***, 1423–1425.

Santure, AW, Stapley, J, Ball, AD, Birkhead, TR, … Slate, J. (2010). On the use of large marker panels to estimate inbreeding and relatedness: empirical and simulation studies of a pedigreed zebra finch population typed at 771 SNPs. *Molecular Ecology*, ***19***, 1439–1451.

Schwabl, P, Llewellyn, MS, Landguth, EL, Andersson, B, … Grijalva, MJ. (2017). Prediction and prevention of parasitic diseases using a landscape genomics framework. *Trends in Parasitology*, ***33***, 264–275.

Segelbacher, G, Cushman, SA, Epperson, BK, Fortin, MJ, … Manel, S. (2010). Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation Genetics*, ***11***, 375–385.

Serieys, LE, Lea, A, Pollinger, JP, Riley, SP, & Wayne, RK. (2015). Disease and freeways drive genetic change in urban bobcat populations. *Evolutionary Applications*, ***8***, 75–92.

Shirk, AJ, Cushman, SA, & Landguth, EL. (2012). Simulating pattern-process relationships to validate landscape genetic models. *International Journal of Ecology*, ***2012***, e539109.

Shirk, AJ, Landguth, EL, & Cushman, SA. (2018). A comparison of regression methods for model selection in individual-based landscape genetic analysis. *Molecular Ecology Resources*, ***18***, 55–67.

Shochat, E, Lerman, SB, Anderies, JM, Warren, PS, … Nilon, CH. (2010). Invasion, competition, and biodiversity loss in urban ecosystems. *BioScience*, ***60***, 199–208.

Short Bull, RA, Cushman, SA, MacE, R, Chilton, T, … Luikart, G. (2011). Why replication is important in landscape genetics: American black bear in the Rocky Mountains. *Molecular Ecology*, ***20***, 1092–1107.

Slate, D, Algeo, TP, Nelson, KM, Chipman, RB, … Rupprecht, CE. (2009). Oral rabies vaccination in North America: opportunities, complexities, and challenges. *PLoS Neglected Tropical Diseases*, ***3***, 1–9.

Sleeman, JM, Keane, JM, Johnson, JS, Brown, RJ, & Woude, S V. (2001). Feline leukemia virus in a captive bobcat. *Journal of Wildlife Diseases*, ***37***, 194–200.

Smith, KF, Acevedo-Whitehouse, K, & Pedersen, AB. (2009). The role of infectious diseases in biological conservation. *Animal Conservation*, ***12***, 1–12.

Smith, KF, Sax, DF, & Lafferty, KD. (2006). Evidence for the role of infectious disease in

species extinction and endangerment. *Conservation Biology*, ***20***, 1349–1357.

Smith, MD, Wertheim, JO, Weaver, S, Murrell, B, … Kosakovsky Pond, SL. (2015). Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Molecular Biology and Evolution*, ***32***, 1342–1353.

Spielman, D, Brook, BW, Briscoe, DA, & Frankham, R. (2004). Does inbreeding and loss of genetic diversity decrease disease resistance? *Conservation Genetics*, ***5***, 439–448.

Storfer, A, Epstein, B, Jones, M, Micheletti, S, … Fox, S. (2017). Landscape genetics of the Tasmanian devil: implications for spread of an infectious cancer. *Conservation Genetics*, ***18***, 1287–1297.

Storfer, A, Murphy, MA, Evans, JS, Goldberg, CS, … Waits, LP. (2007). Putting the "landscape" in landscape genetics. *Heredity*, ***98***, 128–142.

Storfer, A, Murphy, MA, Spear, SF, Holderegger, R, & Waits, LP. (2010). Landscape genetics: where are we now? *Molecular Ecology*, ***19***, 3496–3514.

Storfer, A, Patton, A, & Fraik, AK. (2018). Navigating the interface between landscape genetics and landscape genomics. *Frontiers in Genetics*, ***9***.

Streicker, DG, Winternitz, JC, Satterfield, DA, Condori-Condori, RE, … Valderrama, W. (2016). Host–pathogen evolutionary signatures reveal dynamics and future invasions of vampire bat rabies. *Proceedings of the National Academy of Sciences*, ***113***, 10926–10931.

Symula, RE, Marpuri, I, Bjornson, RD, Okedi, L, … Caccone, A. (2011). Influence of host phylogeographic patterns and incomplete lineage sorting on within-species genetic variability in *Wigglesworthia* species, obligate symbionts of tsetse flies. *Applied and Environmental Microbiology*, ***77***, 8400–8408.

Talbot, B, Garant, D, Paquette, SR, Mainguy, J, & Pelletier, F. (2012). Lack of genetic structure and female-specific effect of dispersal barriers in a rabies vector, the striped skunk (*Mephitis mephitis*). *PLoS One*, ***7***, e49736.

Talbot, B, Vonhof, MJ, Broders, HG, Fenton, B, & Keyghobadi, N. (2017). Comparative analysis of landscape effects on spatial genetic structure of the big brown bat and one of its cimicid ectoparasites. *Ecology and Evolution*, ***7***, 8210–8219.

Taylor, PD, Fahrig, L, Henein, K, & Merriam, G. (1993). Connectivity is a vital element of landscape structure. *Oikos*, ***68***, 571–573.

Tesson, SVM, Okamura, B, Dudaniec, RY, Vyverman, W, … Löndahl, J. (2016). Integrating microorganism and macroorganism dispersal: modes, techniques and challenges with particular focus on co-dispersal. *Ecoscience*, ***22***, 109–124.

Thomassen, HA, Harrigan, RJ, Semple Delaney, K, Riley, SPD, … Smith, TB. (2018). Determining the drivers of population structure in a highly urbanized landscape to inform conservation planning. *Conservation Biology*, ***32***, 148–158.

Tigas, LA, Van Vuren, DH, & Sauvajot, RM. (2002). Behavioral responses of bobcats and coyotes to habitat fragmentation and corridors in an urban environment. *Biological Conservation*, ***108***, 299–306.

Tonteri, A, Vasemägi, A, Lumme, J, & Primmer, CR. (2010). Beyond MHC: Signals of

elevated selection pressure on Atlantic salmon (*Salmo salar*) immune-relevant loci. *Molecular Ecology*, ***19***, 1273–1282.

Torres-Pérez, F, Palma, RE, Hjelle, B, Holmes, EC, & Cook, JA. (2011). Spatial but not temporal co-divergence of a virus and its mammalian host. *Molecular Ecology*, ***20***, 4109–4122.

Townson, H, Nathan, MBM, Zaim, M, Guillet, P, … Kindhauser, M. (2005). Exploiting the potential of vector control for disease prevention. *Bulletin of the World Health Organization*, ***83***, 942–947.

Trovão, NS, Suchard, MA, Baele, G, Gilbert, M, & Lemey, P. (2015). Bayesian inference reveals host-specific contributions to the epidemic expansion of influenza A H5N1. *Molecular Biology and Evolution*, ***32***, 3264–3275.

U.S. Census Bureau. (1900). 1900 Census of Population and Housing. *Population and Housing Unit Counts*.

U.S. Census Bureau. (2010). 2010 Census of Population and Housing. *Population and Housing Unit Counts*.

U.S. Census Bureau. (2016). American Community Survey 1-year estimates.

United Nations. (2018). World Urbanization Prospects. *United Nations Department of Economic & Social Affairs Population Division*.

Van Strien, MJ, Keller, D, & Holderegger, R. (2012). A new analytical approach to landscape genetic modelling: least-cost transect analysis and linear mixed models. *Molecular Ecology*, ***21***, 4010–4023.

Vander Wal, E, Edye, I, Paquet, PC, Coltman, DW, … Andres, JA. (2013). Juxtaposition between host population structures: implications for disease transmission in a sympatric cervid community. *Evolutionary Applications*, ***6***, 1001–1011.

VandeWoude, S, & Apetrei, C. (2006). Going wild: lessons from naturally occurring T-lymphotropic lentiviruses. *Clinical Microbiology Reviews*, ***19***, 728–762.

Volz, EM, Koelle, K, & Bedford, T. (2013). Viral phylodynamics. *PLoS Computational Biology*, ***9***.

Volz, EM, Romero-Severson, E, & Leitner, T. (2017). Phylodynamic inference across epidemic scales. *Molecular Biology and Evolution*, ***34***, 1276–1288.

Wagner, HH, & Fortin, M-J. (2013). A conceptual framework for the spatial analysis of landscape genetic data. *Conservation Genetics*, ***14***, 253–261.

Wang, IJ. (2009). Fine-scale population structure in a desert amphibian: landscape genetics of the black toad (*Bufo exsul*). *Molecular Ecology*, ***18***, 3847–3856.

Wang, IJ. (2010). Recognizing the temporal distinctions between landscape genetics and phylogeography. *Molecular Ecology*, ***19***, 2605–2608.

Wang, IJ, & Bradburd, GS. (2014). Isolation by environment. *Molecular Ecology*, ***23***, 5649–5662.

Waples, RK, Larson, WA, & Waples, RS. (2016). Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of

loci. *Heredity*, ***117***, 233–240.

Wenzel, MA, Douglas, A, James, MC, Redpath, SM, & Piertney, SB. (2016). The role of parasite-driven selection in shaping landscape genomic structure in red grouse (*Lagopus lagopus scotica*). *Molecular Ecology*, ***25***, 324–341.

White, LA, Forester, JD, & Craft, ME. (2017). Using contact networks to explore mechanisms of parasite transmission in wildlife. *Biological Reviews*, ***92***, 389–409.

Wickham, H, & Chang, W. (2008). ggplot2: an implementation of the Grammar of Graphics.

Wright, S. (1943). Isolation by distance. *Genetics*, ***28***, 114–138.

Young, SG, Carrel, M, Kitchen, A, Malanson, GP, … Kayali, G. (2017). How's the flu getting through? Landscape genetics suggests both humans and birds spread H5N1 in Egypt. *Infection, Genetics and Evolution*, ***49***, 293–299.

Ypma, RJF, van Ballegooijen, WM, & Wallinga, J. (2013). Relating phylogenetic trees to transmission trees of infectious disease outbreaks. *Genetics*, ***195***, 1055–1062.

Zeller, KA, Creech, TG, Millette, KL, Crowhurst, RS, … Landguth, EL. (2016). Using simulations to evaluate Mantel-based methods for assessing landscape resistance to gene flow. *Ecology and Evolution*, ***6***, 4115–4128.